Supplementary Information for Tung et al, "Social environment is associated with gene regulatory variation in the rhesus macaque immune system"

1. Supplementary Information Materials and Methods

2. Supplementary Tables

3. Supplementary Figures

1. Supplementary Information Materials and Methods

Study population and samples

Individuals included in this study were members of ten social groups of female rhesus macaques housed at the Yerkes NPRC. Most groups had been formed 5 years previously as a part of studies on the relationships between psychosocial stress, reproduction, metabolism, and behavior (e.g., 1). Two groups were formed more recently following the same protocols (Table S1). However, removing these two groups from the data set produced qualitatively similar results as including them (r = 0.950 for the correlation between rank effects on gene expression, by gene, estimated with and without including these new groups), and prediction accuracy for relative rank class for these females did not differ from that for other groups (p = 0.755). As described previously (2), groups were formed by removing females from the large breeding groups at the Yerkes NPRC Field Station and placing them in separate housing. Prior to new group formation, all females were ovariectomized. Females in the middle part of the dominance hierarchy were selected to ensure that all had a similar social history. Unfamiliar females were randomly introduced sequentially to indoor-outdoor run housing (25 m by 25 m for each area) over the course of one week, until all groups included five adult females. Dominance hierarchies formed quickly with minimal contact aggression. A single male was also housed in each social group at the time these individuals were sampled.

Dominance rank at time of sampling was strongly correlated with order of introduction (Spearman's rho = 0.61, p = 3.25 x 10^{-6}; excluding one group with multiple rank shifts—used in our plasticity analysis—rho = 0.72, p = 2.94 x 10^{-8}; note that because >1 individual is necessary to constitute a "group," the first two individuals added were tied in terms of order of introduction, such that the maximum value of rho <1). In contrast, we identified no significant correlation between dominance rank and age (p = 0.34), parity (p = 0.12), time since oviarectomy (p = 0.16), or time since removal from the large breeding colonies (p = 0.52). Furthermore, we also identified no signal of any of these variables on gene expression after controlling for social group (beyond that expected by chance: see Figure S1).

We chose to assess rank-dependent gene expression in PBMCs both because of the accessibility of this tissue, and because of the relevance of PBMCs to immune function, which has close ties to social stress. To sample PBMCs, we obtained blood samples for 49 of the 50 individuals in these social groups in heparinized Vacutainer tubes, as well as replicate samples from a different sampling effort for 7 of these 49 females (changes in dominance rank in these samples occurred within the course of 1 year). For one group, we sampled only four individuals because the lowest ranking female had recently been removed from the group and had not yet been replaced. All study subjects had previously been habituated to conscious venipuncture, and group members were sampled within 10 minutes of entrance into the housing area. Peripheral blood mononuclear cells (PBMCs) were extracted from these samples using a Ficoll gradient, and RNA (for gene expression analysis) or DNA (for methylation analysis) was purified from the total PBMC fraction using the Qiagen RNEasy kit or the Qiagen Gentra Puregene kit respectively. For cell-type specific expression analyses, we physically separated $CD3^+/CD4^+$ (e.g., helper T cell), $CD3^+/CD8^+$ (e.g., cytotoxic T cell), monocyte ($CD14^+$), and B cell ($CD20^+$) populations prior to RNA extraction from each cell type.

Blood samples for PBMC purification were collected over a period of three weeks (Table S1). During each sampling period, individual ranks were confirmed using focal sampling (3) to record the outcomes of dyadic agonistic interactions, with subordinate status defined when an

animal regularly emitted unequivocal submissive gestures to another animal. Agonistic behaviors were identified using a predefined ethogram. Dominance status was defined by which female submitted to which other females and not by which group member was the most aggressive: the 5th ranking female within a group therefore submitted to all other group members, and the highest ranking (alpha) female submitted to none. However, as shown in Fig. 1a, aggression (largely non-contact threats) received by group mates occurred proportionately more often with lower status. In agreement with patterns of rank stability in large mixed-sex groups, these rank assignments have remained largely stable over the course of observation on these animals, with the exception of the 7 cases we used to test for plasticity of the rank-gene expression relationship

FACS analysis

To assess the relative proportions of the major cell types found in PBMCs, we stained a subsample of the purified PBMCs for 39 individuals with five dye-labeled monoclonal antibodies: anti-CD3$^+$ (T cells: CD3-PE-Cy7, BD 557749), anti-CD4$^+$ (associated with helper T cells: CD4-PerCP-Cy5.5, BD Pharmingen 552838), anti-CD8$^+$ (associated with cytotoxic T cells: CD8-APC, Beckman Coulter IM2469U), anti-CD14$^+$ (monocytes: CD14-FITC, Beckman Coulter IM0645U), and anti-CD20$^+$ (B-cells: CD20-PE, BD Pharmingen 555623). For the other ten individuals in our data set, a FACS machine was unavailable at the time of sampling. To account for differences in antibody staining efficiency between batches of samples, we used as our measurement, for each sample, the proportion of stained live cells of each cell type among the live cells of any type. To control for differences among social groups, we regressed out social group effects and used the residuals in our subsequent analyses.

To establish expectations for cell type-specific gene expression levels, we antibody stained purified PBMCs as described above for five individuals. These samples were collected separately from the samples used in our main analysis of dominance rank and gene expression levels, and were chosen to represent five different social groups and all five possible rank positions. We submitted the total PBMC samples to the University of Chicago Flow Cytometry Facility for physical separation of cell populations. A CD3$^+$/CD4$^+$ cell fraction, a CD3$^+$/CD8$^+$ fraction, a CD14$^+$ fraction, and a CD20$^+$ fraction were sorted on a BD FACSAria. RNA extractions were conducted separately for each individual-cell type combination (n = 20) for downstream gene expression profiling (see below; Fig S3). Two of these samples (one of B cells and one of CD3$^{+/}$CD8$^+$ T cells, both from the same individual) were clear outliers from all other samples and were therefore removed from subsequent analyses.

Illumina HT-12 array cross-hybridization, probe quality control, and general gene expression patterns

To measure gene expression levels, we utilized a human microarray platform, the Illumina HT-12 Expression BeadChip. The genome sequences of rhesus macaques and humans are largely similar, especially in coding regions (~96.5%). The set of probes we utilized contained only few mismatches (median sequence similarity = 48 of 50 base pairs), such that the attenuation of hybridization intensity was minimal. Since our study only involved comparisons of gene expression levels within a species, the sequence mismatches with respect to the array probes are not expected to result in biased estimates of gene expression levels (4). Indeed, we have previously demonstrated that cross-species hybridization to an array designed for a closely related species (and specifically cross-array hybridization involving rhesus macaques and probes based on human DNA sequence) does not produce measurable bias in analyses of differences in

gene expression levels, or greatly reduce power, for *within-species* comparisons (5). We acknowledge that a high-throughput sequencing strategy would have allowed us to capture gene expression variation in greater detail and avoided the need for cross-species hybridization. However, such an approach was cost prohibitive at the time we collected these data.

All RNA samples were hybridized to Illumina HT-12 Expression BeadChips at the University of California Los Angeles Southern California Genotyping Consortium core facility. For the main analysis of rank-gene expression relationships, we conducted technical replicate hybridizations in duplicate (n = 47 individuals) or triplicate (n = 2 individuals). We performed one hybridization per sample for the cell-type specific analyses, based on the very high levels of concordance between technical replicates for our earlier analyses (mean *r* between technical replicates = 0.987, range = 0.978 – 0.995). Our study design resulted in 4 – 5 biological replicates of gene expression measurements per cell type.

For probe quality control, we mapped the 50 base pair sequence for each Illumina HT-12 probe to the rhesus macaque genome (rhemac 2.0) using *blat*. We removed any probe that failed to map to the macaque genome (n = 19,384 of 47,232 initial probes) or that mapped to multiple places (n = 3,728) in the macaque genome at 80% identity or higher (40 of 50 base pairs). Note that, although a large percentage of Illumina HT-12 probes were culled as a result of this procedure, this phenomenon also occurs when mapping Illumina HT-12 probes to the human genome (using the same methods, 10,981 probes fail to map uniquely to the hg18 version of the human genome). We also removed any probes that overlapped with the location of known segmental duplicates in rhesus macaque greater than 1 kb in length (n = 753 probes). These steps resulted in a set of 23,367 probes that uniquely mapped to the rhesus macaque genome outside of duplicated regions. We also eliminated probes that were not significantly detected in any sample at $p < 0.001$ (compared to Illumina HT-12 control probes), resulting in a set of 7,303 probes (representing 6,097 genes) in the final data set. For downstream analysis, the gene expression data were $\log_2$ transformed and quantile normalized between arrays using the R package *lumi* (6).

Principal components analysis on mean-centered gene expression data (after regressing out differences in means across social groups) was performed in R using *prcomp* (*stats* package), with the data scaled to unit variance. In addition to investigating the relationship between dominance rank and the resulting PCs, we also explicitly checked for possible effects of age, parity, time since oviarectomy, and time since removal from the large breeding colonies from which these individuals originated. Among the top 10 PCs, which together account for 60% of the variance in the data, we found that dominance rank also correlates with PC4 (which explains 5.1% of variance in the data; rank-PC4 correlation $p = 0.001$). Time since oviarectomy and time since removal from the breeding colony both correlated with PC8 (p = 0.016 and p = 0.008, respectively), which explained 3.1% of overall variance. These weak or absent effects of non-rank variables are consistent with the absence of evidence for such effects in gene-by-gene analyses (Figure S1) and the absence of detectable correlations between these variables and dominance rank reported above.

Linear effects of rank on gene expression

To assess linear relationships between inter-individual variation in dominance rank and expression levels for each gene, we used the following linear mixed-effects model, which accounts for relatedness in the sample (e.g., (7, 8)):

$$y_{ij} = \beta_j r_i + u_{ij} + b_j + \varepsilon_{ij}$$

Here, y represents the residuals of the normalized, log transformed gene expression levels after controlling for the effect of social group by regressing out mean differences across groups. We utilized these residuals as our measure of gene expression in all subsequent analyses in order to take account of possible biological differences in means across social groups, and to take account of batch effects: with few exceptions, all individuals in a group were sampled at the same time and processed together. Thus, controlling for social group also effectively controlled for sample batch effects, because social groups were subsumed within batches (multiple social groups were sometimes sampled on the same day). Individuals are indexed by i and genes are indexed by j. For gene j, β_j is therefore the fixed effect of rank r_i, b_j is the intercept, and ε_{ij} is the residual error, assumed to be normally distributed with mean zero and variance s^2. The term u_{ij} refers to the random effect component of the model, where Var(u_j) equals the estimated genetic variance in y multiplied by the pairwise kinship matrix, K. Models were fitted using the R package *emma* (7), with minor modifications to the source code to accommodate gene expression data (this code is available on the Gilad lab website: http://giladlab.uchicago.edu/data/Tung_Rcode/). We evaluated the significance of β_j, the rank effect, as evidence for a linear relationship between dominance rank and gene expression (Figure S1). False discovery rates were evaluated using the method of Storey and Tibshirani (9), implemented in the R package *qvalue*.

Because the social groups in this study were artificially constructed, the species-typical pattern of matrilineal rank inheritance in rhesus macaques did not pertain to our sample. However, some individuals in the data set as a whole were related. As genetic effects can also have an impact on gene expression levels, we evaluated pairwise relatedness between all individuals in the sample by genotyping 51 highly polymorphic microsatellite loci (2.32% missing data) and estimating relatedness using the program COANCESTRY (10, 11). Mean pairwise relatedness in the sample was 0.028 (+/- 0.051 s.d.), and estimated relatedness was highly correlated with pedigree-based estimates of relatedness available for 7 known dyads (Pearson's r = 0.899, $p = 0.00588$).

Prediction of rank class using gene expression data

To investigate the predictive value of gene expression data for relative position in the social rank hierarchy, we defined relative ranks (i.e., rank classes) based on prior work, which treated ranks 1 and 2 in 5-female hierarchies as high ranking and ranks 3, 4, and 5 in these hierarchies as low ranking (e.g., 1). We also included an intermediate class based on the agonism data on these study subjects, which suggested additional separation between rank 3 and ranks 4 and 5 (Figure 1; Tukey's HSD test, $p < 0.05$ for the rank 3 – rank 5 contrast and $p < 0.09$ for the rank 3- rank 4 contrast; in contrast, $p = 0.995$ for the rank 4 – rank 5 contrast). We refer to these classes as class A (high), class B (middle), and class C (low).

For each iteration of our prediction analysis, we divided the data set into a training set of 39 individuals (~80% of the data) and a test set of 10 individuals (~20% of the data). Each test set contained 2 randomly chosen individuals of each rank (results were highly similar under completely random assignment that was not balanced across ranks: Figure S2). We used the support vector machine (SVM) approach implemented in the program *svm-multiclass* (12) to develop a model that related the combined gene expression data across all measured genes to rank class. The SVM model was fit using the training set data only, with gene expression values rescaled from 0 to 1 for stability. We then assessed model predictive ability as the percentage of the test set individuals for whom rank class was correctly assigned by the fitted model. We also performed a complementary assessment of model predictive ability by calculating the absolute

error in model prediction (the sum of the absolute value of the true rank class minus the predicted rank class, across all test set individuals: each classification mistake of individual of true rank class A into rank class C, or vice versa, added a value of 2 errors, while each classification mistake of individual of true rank class A or C into rank class B, or vice-versa, added a value of 1 error). We compared the median sum of absolute errors in our cross-validation iterations to the distribution of sum of absolute errors under random assignment of rank class. Note that although we report the results of a three-class classification scheme in the main text, retaining a two-class division (high versus low) yielded as good or better prediction accuracy (average leave-k-out accuracy within the main sample set = 94.7%; prediction accuracy across temporally separated samples = 85%).

We also tested whether gene expression data predicted rank class across time, in individuals who changed their ordinal rank positions. For prediction across temporal replicates, we restricted our analysis to probes represented on both the HT-12 v. 3.0 array (used for the seven temporal replicates) and the HT-12 v. 4.0 array (used to run all other arrays in the study). We normalized, log-transformed, and removed the effects of social group for this common set of probes only (6,735 probes with detectable signal in our data set versus 7,303 such probes in the main analysis). We then treated the main data set of 49 individuals as a training set and the expression data set for the 7 individuals when at an earlier rank as a test set. Model accuracy was calculated as the percentage of the unlabeled test set individuals for whom relative rank position was correctly assigned by the fitted model. Note that in the main text, we report the results of prediction across temporal replicates after normalizing and scaling (from 0 to 1 for each gene expression measurement, across samples) the training set and test set gene expression data together. An alternative approach is to normalize and scale each data set independently. This approach also resulted in excellent prediction accuracy (100% of test set samples were classified into the correct rank class, versus 85.7% accuracy achieved through normalizing all samples together; we report the more conservative estimate in the text).

Effects of cell type proportion

To estimate the expected expression level of each gene based on the cell type composition of each sample, we first measured the expression level for each gene in pure populations of each of four cell types (helper T cells, cytotoxic T cells, monocytes, and B cells, obtained via FACS sorting described above). Pure populations were collected from five individuals across all ranks and five different social groups, and we considered the median expression value across individuals as the estimated expression level for each gene in a given cell type. For the 39 samples for which we were able to obtain cell type composition data, we then weighted the estimated cell type-specific expression levels for each gene by the proportion of the appropriate cell type in the PBMC pool for each sample. This value corresponded to the expected gene expression level in the PBMC sample from that individual, if cell type composition represented the sole mechanism underlying variation in gene expression.

We used these values to investigate the contribution of tissue composition to the rank-gene expression relationship identified for the 987 rank responsive genes, using a partial correlations approach. We reasoned that, if tissue composition effects significantly contributed to the rank-gene expression relationship, then the magnitude of the rank-gene expression partial correlation should decrease when conditioned on the expected expression level of that gene based on the cell type composition data. The magnitude of this decrease should also be larger than that calculated when the rank-gene expression partial correlation was conditioned on

permuted values of these expected gene expression levels. Hence, cases in which the rank-gene expression relation was likely explained by tissue composition exhibited:

$$\| Cor(r, e \mid c) \| < \| Cor(r, e \mid c_{permuted}) \|$$

where r represents dominance rank; e represents the residuals of the normalized, log transformed gene expression levels after controlling for the effect of social group; and c represents the expected expression level of a gene based on cell type composition alone.

We conducted 1000 permutations of $c_{permuted}$ in order to establish a null distribution. In cases in which the true rank-gene expression partial correlation fell within this null distribution, we failed to reject the null hypothesis (that tissue composition did not significantly contribute to the rank-gene expression relationship). In cases in which the true partial correlation was smaller than expected by chance based on the null distribution, we interpreted our results as supportive of a rank-gene expression relationship mediated, at least in part, by a tissue composition effect. Because we viewed such an effect as mechanistically conservative (i.e., it does not require additional changes in gene regulation), we set our p-value threshold at a nominal value of 0.05.

Effects of glucocorticoid regulation

To investigate the relationship between dominance rank and glucocorticoid regulation, we used data from a dexamethasone (Dex; a synthetic glucocorticoid) suppression test, which assesses the state of GC negative feedback. Prior to Dex administration, serum samples were obtained from each female at 1100 hours on the day of the assay. At 1730 hours, females received an injection of Dex (0.25 mg/kg) and samples were obtained at 1100 hours the following morning to assay cortisol. The degree of GC negative feedback was assessed by the change in cortisol levels between the post Dex-treatment sample and the control sample. Cortisol assays were performed in the Yerkes NPRC Biomarkers Core Lab using established procedures. Serum levels of cortisol were determined by radioimmunoassay (RIA) with a commercially available kit (Beckman-Coulter/DSL, Webster TX) previously validated for rhesus macaques. Using 25 µl of serum, the assay has a range from 0.5 to 60 µg/dl with an inter- and intra-assay CV of 4.9% and 8.7%, respectively. To test the relationship between rank and degree of dexamethasone suppression (Dex resistance), we assessed the significance of the rank term in a linear model relating rank to the change in cortisol levels after Dex administration.

To evaluate the contribution of GC regulatory state to the variation in the expression levels of the 987 rank-responsive genes, we again employed a partial correlations approach. This analysis paralleled the analysis we conducted for tissue composition effects, except that the relationship we evaluated was:

$$\| Cor(r, e \mid g) \| < \| Cor(r, e \mid g_{permuted}) \|$$

where r represents dominance rank and e represents gene expression levels, as before; and g represents the residuals of data from the Dex clearance trials after controlling for social group effects. As before, we established a null distribution using 1000 permutations of $g_{permuted}$. In cases in which the true rank-gene expression partial correlation fell within this null distribution, we failed to reject the null hypothesis (that glucocorticoid resistance did not significantly contribute to the rank-gene expression relationship). Conversely, in cases in which the true partial correlation was smaller than expected by chance from the null distribution, we interpreted our results as supportive of a rank-gene expression relationship mediated, at least in part, by a GC regulatory effect. Because such an effect is conservative with respect to prior knowledge about the physiology of dominance rank effects, and does not require additional changes in gene regulatory mechanisms, we again set our p-value threshold to a relatively relaxed nominal value

of 0.05. Note however that, unlike the case of for tissue composition, we were not able to estimate gene expression levels under a scenario in which GC-mediated effects are the sole contributor. Thus, it is possible that for some genes identified via this method, the rank-gene expression relationship and the rank-GC signaling relationship are statistically correlated but mechanistically independent.

Joint analysis of tissue composition effects and GC effects were pursued in a parallel manner, except that we evaluated whether:

$$\| Cor(r, e \mid c, g) \| < \| Cor(r, e \mid c_{permuted}, g_{permuted}) \|$$

Bisulfite sequencing and low level data processing for DNA methylation

To measure DNA methylation levels via whole genome bisulfite sequencing, we prepared four sequencing libraries for each individual from DNA obtained from purified PBMCs, following the method of Lister et al (13). Unmethylated lambda phage DNA was incorporated into each library in order to assess the efficiency of bisulfite conversion. We then sequenced 8 Illumina HiSeq flow cell lanes for each individual, with each library represented on 2 lanes to minimize PCR duplicates (see Figure S6; with the exception of 2 individuals, for whom low coverage and a truncated sequencing run motivated us to conduct four additional lanes of sequencing). All sequencing runs were single-ended and 75 base pairs in length. We truncated the resulting reads to 70 base pairs to remove lower quality bases near the end of the reads, and used the tool *cutadapt* (14) to trim off any remaining adapter sequence incorporated in the read.

The resulting reads were mapped to a combined rhesus macaque genome (rheMac 2.0) and lambda phage genome using *bismark* (15), which uses a fully bisulfite converted reference genome sequence as the basis for read mapping (Table S2). Total coverage for each CpG site and the number of reads for each site that were methylated were evaluated using the *methylation_extractor* tool in *bismark.* In doing so, we masked the first three base pairs of each aligned read based on evidence that the methylation estimates for these positions were systematically biased (Fig S5, see also (16) for description of a similar phenomenon).

Analysis of rank-related DNA methylation levels

For the neighbor-joining and hierarchical clustering analyses of DNA methylation data, we identified the locations of Ensembl-annotated transcription start sites for the set of 987 rank-responsive genes (based on the gene expression data). Annotated TSSs were available for 811 of these genes, and we randomly chose one TSS per gene if multiple TSS were associated with the same gene. We then calculated methylation levels for each CpG site within 20 kb of each of these TSS, based on the raw percentage of reads for that site with evidence for a methylated CpG. We excluded sites with 0 coverage in any individual. Over all sites, this procedure produced a data matrix of 445,059 CpG methylation levels for each of the six individuals in the sample. We used these data for the neighbor-joining (conducted using the R package *ape*) and hierarchical clustering analyses depicted in Figure 5, after regressing out social group from the data for each site.

To assess whether features based on differential methylation data could predict rank-associated differential expression, we used leave-one-out cross-validation on a data set consisting of transcripts from the 987 rank-responsive genes and transcripts from the 1,000 genes with the least evidence for rank-associated variation in expression levels (i.e., largest p-values in the initial analysis of the rank-gene expression relationship). For each leave-one-out iteration, we fit a model relating features based on differential methylation (Table S3) to a binary differential

expression variable (rank-associated genes were assigned a value of 1 and rank-independent genes were assigned a value of -1) using *svm-perf* (12, 17). Overall prediction accuracy was measured as the percent of "left-out" genes (of 1,987 total genes) for which the gene expression state was correctly predicted. We evaluate significance of this prediction accuracy by comparison to prediction accuracies calculated in the same way, but on our data set with the labels (rank-associated or rank-independent) permuted, across 100 permutations.

To identify rankDMRs, we used the *bsmooth* method implemented in the R package *bsseq*. Briefly, this approach uses windows of 70 CpG sites or 1,000 base pairs (whichever is larger) to estimate the binomial probability that a CpG site is methylated, smoothing these estimates across nearby CpG sites. This method allowed us to take account of spatial correlations between nearby CpG sites and weight CpG sites with greater coverage more heavily. We limited our analysis to CpG sites with at least 2x coverage in each of the six individuals. We then analyzed the set of CpG t-statistics produced by comparing the smoothed probabilities of methylation for low ranking individuals to the same values for high ranking individuals. We empirically classified differentially methylated regions as regions in which (*i*) contiguous CpGs were no further than 300 base pairs away from each other, (*ii*) all t-statistics were in the extreme end (0.5%) of the genome wide distribution of t-statistic magnitudes, (*iii*) all t-statistics were in the same direction (consistently more methylated in high ranking individuals, or consistently more methylated in low ranking individuals), and (*iv*) at least 3 CpG sites were covered and exhibited at least a 10% difference in mean methylation levels between the two classes. DMRs were merged if they were called within 1 kilobase of each other, with no intervening analyzed CpGs.

To test for overrepresentation of rankDMRs near rank-associated differentially expressed genes, we used *bedtools* (18) to identify the TSS closest to each DMR, within 20 kb or less intervening distance (our results are consistent if we use a smaller cutoff of 15 kb or a larger cutoff of 30 kb: $p < 0.03$ for 30 kb and $p < 0.05$ for 15 kb; smaller distances greatly reduce the sample size of genes). This procedure assigned one transcript (gene) to each DMR, provided that the DMR was close to a genic region. Within this set, we counted the number of appearances of rank-associated differentially expressed genes (n = 164 TSS, representing 162 discrete genes) and the number of appearances of non-differentially expressed genes (n = 654 TSS). We used a two-tailed Fisher's exact test to compare these numbers to the numbers of cases of differentially expressed (1,351 TSS) and non-differentially expressed (6,577 TSS) genes in general. In all cases in which multiple transcripts for a gene utilized the same TSS, we counted that TSS only once.

2. Supplementary Tables

Table S1. Study subjects[1].

Animal ID	Social group	Date of introduction (month/year)	Rank (at time of sampling)	Sample date (month/year)	Time in group (months)
RBm4	1	07/05	1	07/10	60
ROh4	1	07/05	2	07/10	60
RHn6	1	07/05	3	07/10	60
RMg5	1	10/08	4	07/10	21
RVi4	1	07/05	5	07/10	60
RBe5	2	07/05	1	08/10	61
RHc4	2	07/05	2	07/10	60
RMg3	2	07/05	3	07/10	60
RRb7	2	07/05	4	07/10	60
RZt5	2	07/05	5	07/10	60
RZr2	3	03/10	1	07/10	4
RVh5	3	03/10	2	07/10	4
RNu7	3	04/10	3	07/10	3
RBk7	3	08/10	4	08/10	<1
RGv6	4	07/05	1	07/10	60
RTr4	4	07/05	2	07/10	60
RCk4	4	07/05	3	07/10	60
RWe7	4	07/05	4	07/10	60
RCt4	4	07/10	5	08/10	1
RZp6	5	07/05	1	07/10	60
RIz6	5	07/05	2	07/10	60
RYn5	5	07/05	3	07/10	60
RRu6	5	07/05	4	07/10	60
RZd7	5	07/05	5	07/10	60
RWu4	6	07/05	1	08/10	61
RCv6	6	07/05	3	08/10	61
RDv6	6	07/05	4	08/10	61
RId7	6	07/05	2	08/10	61
RJc6	6	07/10	5	08/10	61
RNf6	7	07/05	1	08/10	61
RZk6	7	07/05	2	08/10	61
RQq4	7	07/05	3	08/10	61
RFc6	7	07/05	4	08/10	61
RNf4	7	12/08	5	08/10	20
REm6	8	07/05	1	08/10	61
RTv6	8	07/05	2	08/10	61
ROb6	8	07/05	4	08/10	61
RRa7	8	07/05	3	08/10	61
RGs6	8	07/05	5	08/10	61
ROy4	9	07/05	1	08/10	61
RYh4	9	07/05	2	08/10	61
RWb7	9	07/05	3	08/10	61
RFp8	9	02/10	4	08/10	6
RIp7	9	07/09	5	08/10	13
RMu3	10	03/10	1	08/10	5
RVg5	10	03/10	2	08/10	5
RDe3	10	03/10	3	08/10	5
RUo4	10	03/10	5	08/10	5
RMf4	10	03/10	4	08/10	5

[1]Individuals highlighted in red were those sampled at two different times while occupying 2 different ranks. At the alternative time point, RRa7 was ranked 4, RWb7 was ranked 5, RVi4 was ranked 2, RBm4 was ranked 4, ROh4 was ranked 1, RId7 was ranked 4, and RMg5 was ranked 5.

Table S2. Summary of raw bisulfite sequencing data.

Individual	Rank	Total number of reads[1]	Total uniquely mapped reads[2]	Number of spike-in reads[3]	Conversion efficiency	Mean CpG coverage
REm6	1	7.24e8	4.39e8	2.99e5	99.6%	11.30
RGs6	5	7.95e8	4.61e8	3.28e5	99.6%	11.78
RNf4	5	6.57e8	4.41e8	3.27e5	99.5%	13.91
RNf6	1	6.04e8	3.91e8	3.22e5	99.6%	11.80
RZd7	5	6.36e8	4.11e8	3.37e5	99.5%	14.05
RZp6	1	5.76e8	3.58e8	2.91e5	99.6%	10.90

[1]For REm6 and RGs6, 2.00e8 and 2.18e8 reads were 50 bp reads instead of 75 (truncated to 70 base pairs for mapping purposes) bp reads

[2]With apparent PCR duplicates removed

[3]Reads mapped to cl857 lambda DNA after filtering for PCR duplicates and removing non-uniquely mapped reads

Table S3. Features used for SVM classification of rank-associated or rank-independent gene expression using DNA methylation data.

1	distance to nearest DMR
2	log10 distance (in base pairs) to the nearest DMR
3	t-statistic area (sum of all t-statistics for CpGs in a DMR) for the nearest dmr
4	number of CpG sites covered by the nearest dmr
5	width of the nearest DMR in base pairs
6	t-statistic area x distance (bps) to the nearest DMR
7	t-statistic area x log10 distance (bps) to the nearest DMR
8	mean t-statistic across all CpGs 1kb upstream of the TSS
9	mean t-statistic across all CpGs 5kb upstream of the TSS
10	mean t-statistic across all CpGs 20kb upstream of the TSS
11	mean t-statistic across all CpGs 50kb upstream of the TSS
12	maximum absolute value of the t-statistic 1kb upstream of the TSS
13	maximum absolute value of the t-statistic 5kb upstream of the TSS
14	maximum absolute value of the t-statistic 20kb upstream of the TSS
15	maximum absolute value of the t-statistic 50kb upstream of the TSS
16	mean t-statistic across all CpGs 1kb downstream of the TSS
17	mean t-statistic across all CpGs 5kb downstream of the TSS
18	mean t-statistic across all CpGs 20kb downstream of the TSS
19	mean t-statistic across all CpGs 50kb downstream of the TSS
20	maximum absolute value of the t-statistic 1kb downstream of the TSS
21	maximum absolute value of the t-statistic 5kb downstream of the TSS
22	maximum absolute value of the t-statistic 20kb downstream of the TSS
23	maximum absolute value of the t-statistic 50kb downstream of the TSS
24	number of CpGs with t-stats in the majority direction, 1 kb upstream of the TSS
25	% of CpGs in the majority direction, 1 kb upstream of the TSS
26	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 1 kb upstream of the TSS
27	% of t-stats for which abs(t-stat) > 2.97, 1 kb upstream of the TSS
28	number of CpGs with t-stats in the majority direction, 5 kb upstream of the TSS
29	% of CpGs in the majority direction, 5 kb upstream of the TSS
30	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 5 kb upstream of the TSS
31	% of t-stats for which abs(t-stat) > 2.97, 5 kb upstream of the TSS
32	number of CpGs with t-stats in the majority direction, 20 kb upstream of the TSS
33	% of CpGs in the majority direction, 20 kb upstream of the TSS
34	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 20 kb upstream of the TSS
35	% of t-stats for which abs(t-stat) > 2.97, 20 kb upstream of the TSS
36	number of CpGs with t-stats in the majority direction, 50 kb upstream of the TSS
37	% of CpGs in the majority direction, 50 kb upstream of the TSS
38	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 50 kb upstream of the TSS
39	% of t-stats for which abs(t-stat) > 2.97, 50 kb upstream of the TSS
40	% of CpGs in the majority direction, 1 kb downstream of the TSS
41	1kb num t stats > abs(2.97) downstream

42	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 1 kb downstream of the TSS
43	% of t-stats for which abs(t-stat) > 2.97, 1 kb downstream of the TSS
44	number of CpGs with t-stats in the majority direction, 5 kb downstream of the TSS
45	% of CpGs in the majority direction, 5 kb downstream of the TSS
46	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 5 kb downstream of the TSS
47	% of t-stats for which abs(t-stat) > 2.97, 5 kb downstream of the TSS
48	number of CpGs with t-stats in the majority direction, 20 kb downstream of the TSS
49	% of CpGs in the majority direction, 20 kb downstream of the TSS
50	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 20 kb downstream of the TSS
51	% of t-stats for which abs(t-stat) > 2.97, 20 kb downstream of the TSS
52	number of CpGs with t-stats in the majority direction, 50 kb downstream of the TSS
53	% of CpGs in the majority direction, 50 kb downstream of the TSS
54	number of t-stats for which abs(t-stat) > 2.97 (0.5% of the genome-wide distribution), 50 kb downstream of the TSS
55	% of t-stats for which abs(t-stat) > 2.97, 50 kb downstream of the TSS
56	total number of t-stats for which abs(t-stat) > 2.97, within 1 kb on either side of the TSS
57	total number of t-stats for which abs(t-stat) > 2.97, within 5 kb on either side of the TSS
58	total number of t-stats for which abs(t-stat) > 2.97, within 20 kb on either side of the TSS
59	total number of t-stats for which abs(t-stat) > 2.97, within 50 kb on either side of the TSS

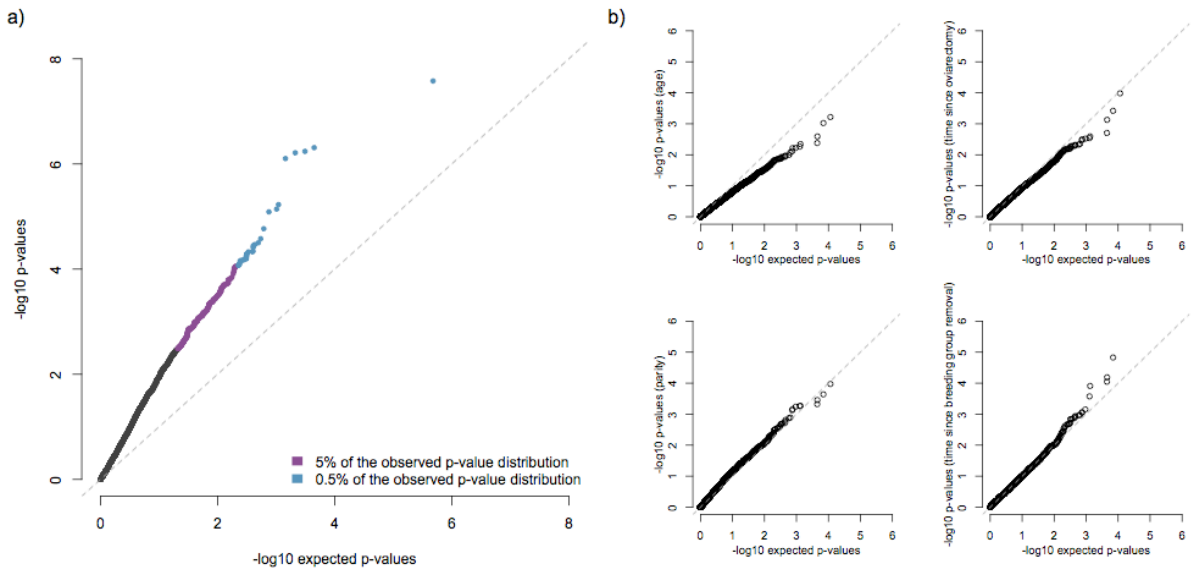

Figure S1. Quantile-quantile plot of p values. a) Quantile distribution of p values for the relationship between dominance rank and gene expression levels in the 6,097 genes we analyzed in our main data (y-axis) compared to quantiles from theoretical p values from a uniform distribution (x-axis); b) Q-Q plots of p values for possible confounding variables (age, parity, time since oviarectomy, time since removal from breeding groups) analyzed using a linear mixed effects model parallel to that run for dominance rank. None of these variables exhibits an enrichment for low p-values, indicating that they do not explain substantial variance in the gene expression data set.

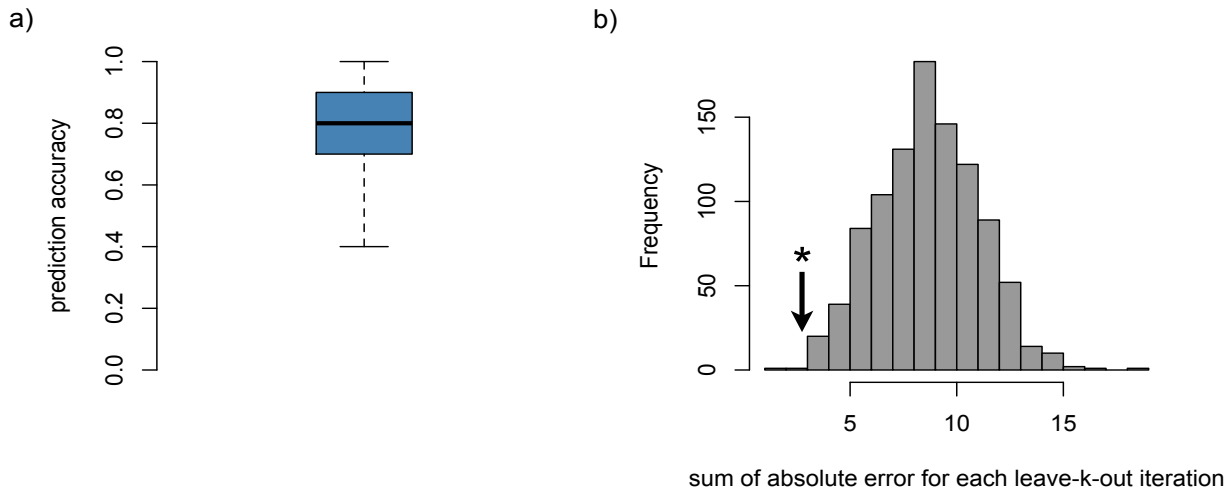

Figure S2. Leave-k-out prediction results, under random assignment of test set/training set membership instead of random assignment with removal of an equal number of individuals of each rank (as reported in the main text). a) boxplot of predictive accuracy for 10 training set individuals obtained across 1000 leave-k-out iterations followed by cross-validation, and b) histogram of the sum of absolute error between predicted rank class and true rank class, if rank classes were randomly assigned. The black arrow and asterisk show the median sum of absolute error across the 1000 true leave-k-out iterations ($p = 0.011$).

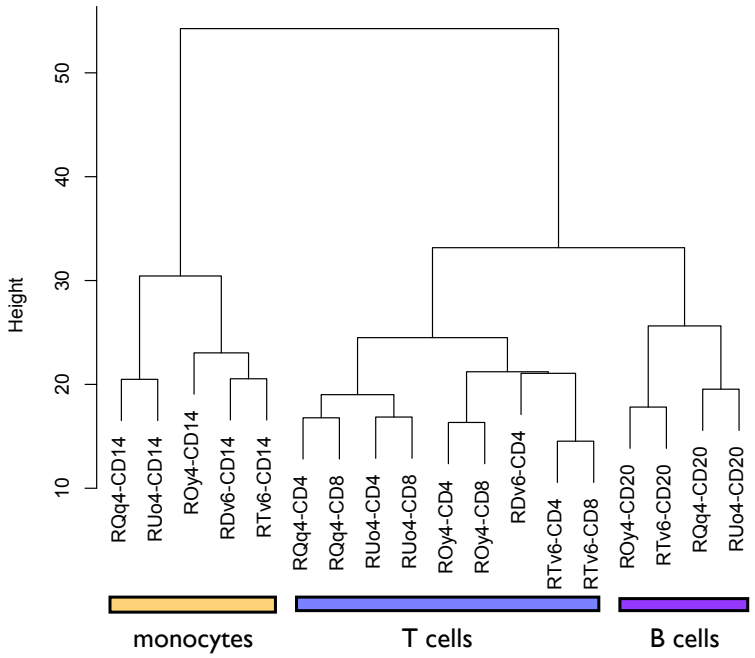

Figure S3. Hierarchical clustering relationships for cell type-specific gene expression profiles. T cells and B cells cluster together, with all monocyte cell fractions outside of this aggregate lymphocyte cluster. Gene expression profiles are more similar within cell types than within individuals, with the exception of $CD4^+$ and $CD8^+$ T cells, which group together within individuals.

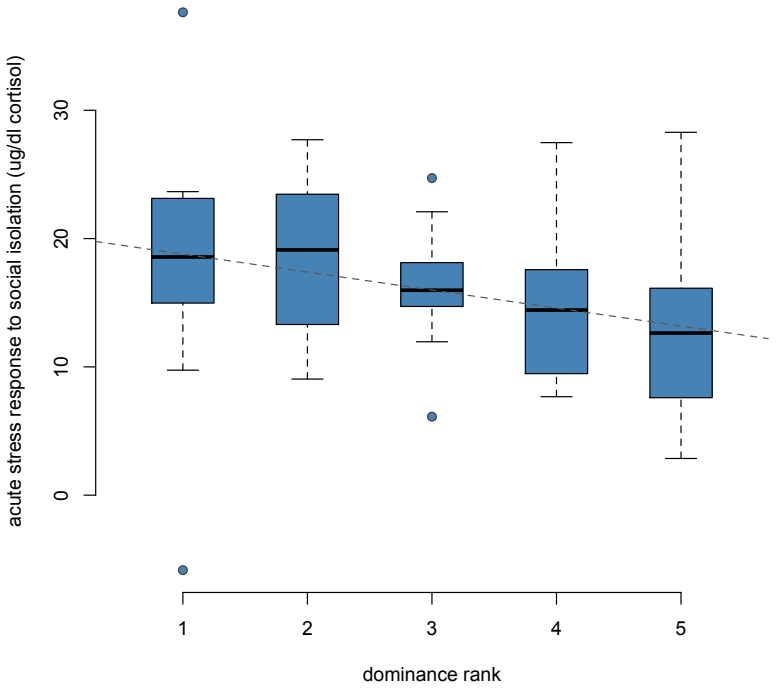

Figure S4. The acute stress response to social isolation tends to be less pronounced in low-ranking individuals. Y-axis depicts the change in cortisol levels between baseline and after social isolation of each individual. Gray line depicts the intercept and slope of a linear regression of this change on dominance rank ($p = 0.067$, $R^2 = 0.070$ $n = 49$).

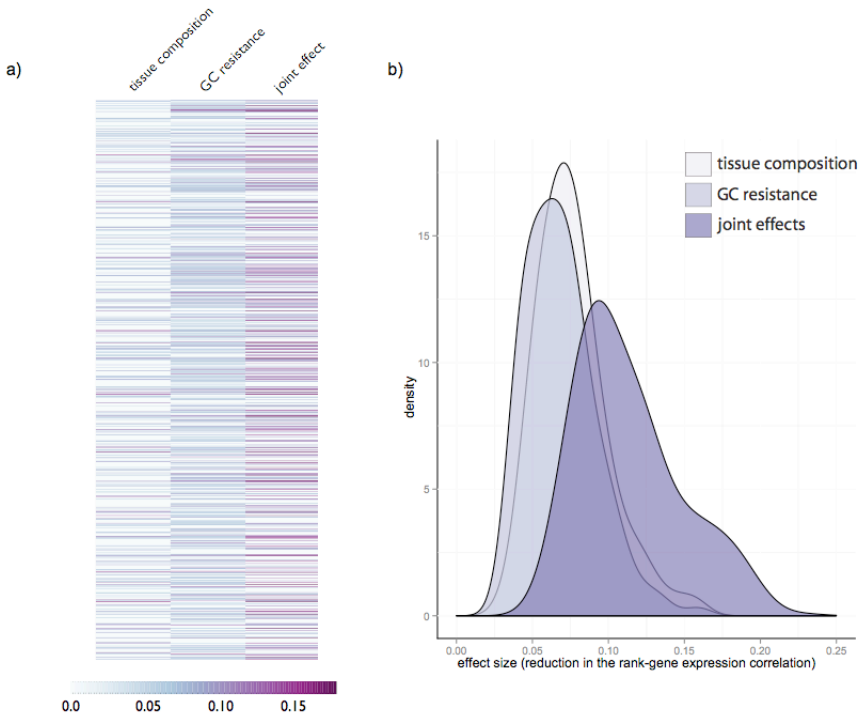

Figure S5. Effect size estimates for putative regulatory mechanisms. a) Reduction in the strength of the dominance rank-gene expression partial correlation when conditioning on tissue composition data (expected gene expression levels based on expression of each gene in each of four individual PBMC cell types and on proportions of these cell types per individual), GC resistance data, or both tissue composition and GC resistance data jointly (e.g., *Cor(rank, gene expression | tissue composition, GC resistance)*). Each line represents a gene for which a significant rank association was detected (n = 987); genes are ordered alphabetically from top to bottom. Darker colors represent larger effect sizes and values for genes for which no significant effect of the mechanism was detected are zeroed out. All effect sizes were calculated by comparing the median value of the dominance rank-gene expression partial correlation, conditioned on permuted data, to the value of the rank-gene expression partial correlation, conditioned on the candidate mechanism(s). b) Distribution of effect sizes when conditioning the rank-gene expression relationship on tissue composition effects, GC resistance data, or both mechanisms jointly.

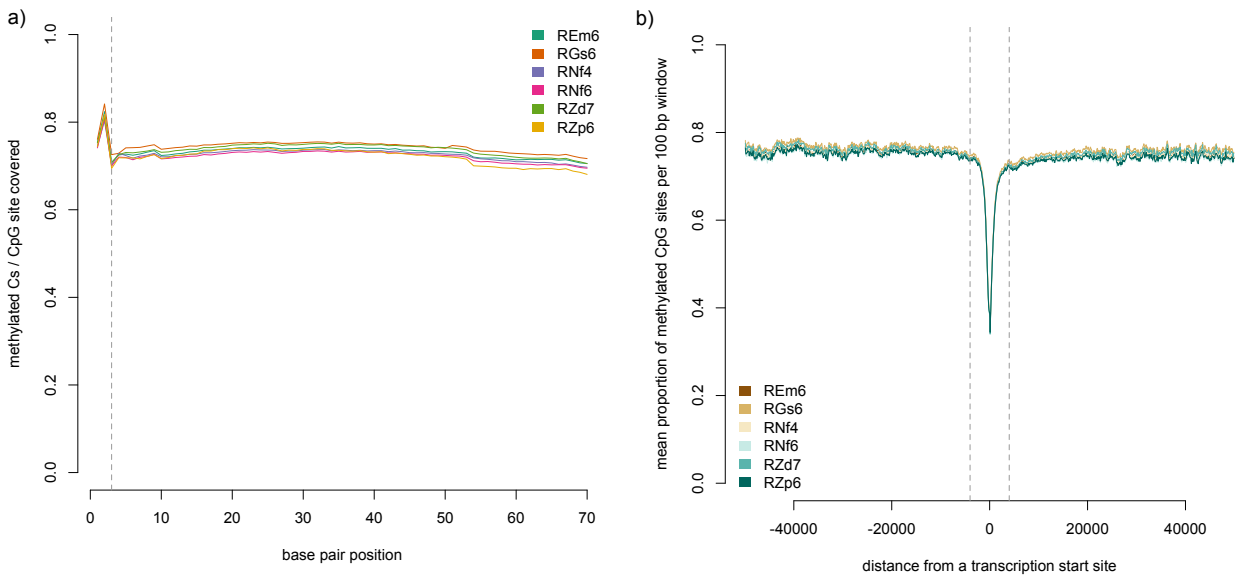

Figure S6. Proportion of CpG sites classified as methylated, by a) base pair position in the sequencing read. The proportion of methylated CpGs observed at each position is expected to be uniform across the length of the read. We observed a systematic bias in DNA methylation estimates for the first three base pair positions in the read, and therefore masked CpG data from these sites when estimating CpG coverage and methylation; and b) distance from annotated transcription start sites. As expected, mean methylation levels drop close to TSS, for all individuals.

References

1. Michopoulos V, Checchi M, Sharpe D, & Wilson ME (2011) Estradiol effects on behavior and serum oxytocin are modified by social status and polymorphisms in the serotonin transporter gene in female rhesus monkeys. *Horm Behav* 59:528-535.
2. Jarrell H*, et al.* (2008) Polymorphisms in the serotonin reuptake transporter gene modify the consequences of social status on metabolic health in female rhesus monkeys. *Physiol Behav* 93:807-819.
3. Altmann J (1974) Observational study of behavior: sampling methods. *Behaviour* 49:227-267.
4. Gilad Y, Rifkin SA, Bertone P, Gerstein M, & White KP (2005) Multi-species microarrays reveal the effect of sequence divergence on gene expression profiles. *Genome Res* 15:674-680.
5. Oshlack A, Chabot AE, Smyth GK, & Gilad Y (2007) Using DNA microarrays to study gene expression in closely related species. *Bioinformatics* 23:1235-1242.
6. Du P, Kibbe WA, & Lin SM (2008) lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24:1547-1548.
7. Kang HM*, et al.* (2008) Efficient control of population structure in model organism association mapping. *Genetics* 178:1709-1723.
8. Yu J*, et al.* (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38:203-208.
9. Storey JD & Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100:9440-9445.
10. Milligan BG (2003) Maximum-likelihood estimation of relatedness. *Genetics* 163:1153-1167.
11. Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol Ecol Resour* 11:141-145.
12. Joachims T (1999) Making large-scale SVM learning practical. *Advances in Kernel Methods - Support Vector Learning*, eds Scholkopf B, Burges C, & Smola A (MIT Press, Boston).
13. Lister R*, et al.* (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462:315-322.
14. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17:10-12.
15. Krueger F & Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27:1571-1572.
16. Hansen KD*, et al.* (2011) Increased methylation variation in epigenetic domains across cancer types. *Nat Genet* 43:768-775.
17. Joachims T (2006) Training linear SVMs in linear time. in *ACM Conference on Knowledge Discovery and Data Mining* (Philadelphia, PA).
18. Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-842.