

Transcriptional modulation of the developing immune system by early life social adversity

Steven W. Cole^{a,1}, Gabriella Conti^b, Jesusa M. G. Arevalo^a, Angela M. Ruggiero^c, James J. Heckman^{d,e,f,1}, and Stephen J. Suomi^f

^aDepartment of Medicine, Division of Hematology-Oncology, and Department of Psychiatry and Biobehavioral Sciences, University of California School of Medicine, Los Angeles, CA 90095; ^bHarris School of Public Policy Studies, University of Chicago, Chicago, IL 60637; ^cLaboratory of Comparative Ethology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892-7971; ^dDepartment of Economics, University of Chicago, Chicago, IL 60637; ^eAmerican Bar Foundation, Chicago, IL 60611; and ^fGeary Institute, University College Dublin, Dublin 4, Ireland

Contributed by James J. Heckman, October 22, 2012 (sent for review March 28, 2012)

To identify molecular mechanisms by which early life social conditions might influence adult risk of disease in rhesus macaques (*Macaca mulatta*), we analyze changes in basal leukocyte gene expression profiles in 4-mo-old animals reared under adverse social conditions. Compared with the basal condition of maternal rearing (MR), leukocytes from peer-reared (PR) animals and PR animals provided with an inanimate surrogate mother (surrogate/peer reared, SPR) show enhanced expression of genes involved in inflammation, cytokine signaling, and T-lymphocyte activation, and suppression of genes involved in several innate antimicrobial defenses including type I interferon (IFN) antiviral responses. Promoter-based bioinformatic analyses implicate increased activity of CREB and NF- κ B transcription factors and decreased activity of IFN response factors (IRFs) in structuring the observed differences in gene expression. Transcript origin analyses identify monocytes and CD4⁺ T lymphocytes as primary cellular mediators of transcriptional up-regulation and B lymphocytes as major sources of down-regulated genes. These findings show that adverse social conditions can become embedded within the basal transcriptome of primate immune cells within the first 4 mo of life, and they implicate sympathetic nervous system-linked transcription control pathways as candidate mediators of those effects and potential targets for health-protective intervention.

stress | social genomics | gene regulation

Exposure to adverse social environments during early life is associated with increased risk of disease in adulthood (1–5), but the biological mechanisms producing such effects remain poorly understood. One possible explanation suggests that neural and endocrine responses to adversity in childhood affect the development of health-relevant molecular systems (i.e., a “defensive programming” of the developing body) (4, 6–10), rendering the body more vulnerable to subsequent pathogen challenges in adulthood (11, 12). Given the transience of most neuroendocrine responses, however, it remains unclear how the extraorganismic social conditions that do “get into the body” during early life could “stay there” over decades to impact the risk of disease in adulthood (13).

One molecular mechanism that could potentially create a persisting biological impact of early life socio-environmental conditions involves the complex systems behavior of the gene transcriptional networks that govern cell growth, differentiation, and function (14, 15). Gene regulatory networks show dynamic landscapes in which the system’s responses to external perturbations converge on a small number of stable “attractor” modes that can subsequently self-perpetuate (16). These self-perpetuating dynamics are sustained in part by the fact that the mRNA “output” of the system at one point in time (i.e., the genome-wide transcriptional profile) constitutes an “input” to the system at subsequent time points because translated mRNA shapes the cell’s response to future environments (17). Mathematical models of human development that capture such recursive dynamics show that small exogenous influences early in life can significantly alter the course of subsequent life trajectories (18–20). What is not

known is which specific genes might be sensitive to such early life environments.

Several recent studies have linked adverse social conditions in early life with adult differences in gene expression in cells of the nervous and immune systems (8–10, 21, 22). Early life social adversity has also been associated with adult cell differences in transcription-related epigenetic features such as DNA methylation (23–25). To determine whether these adult transcriptional alterations might potentially stem from a biological reprogramming of the developing immune system during early life, we analyze the genome-wide transcriptional profile of circulating leukocytes in infant rhesus macaques (*Macaca mulatta*) after 4 mo of experimentally imposed social adversity (peer vs. maternal rearing) (26). To the extent that adverse social conditions become rapidly embedded into the gene regulatory regime of the developing immune system, we expect that (i) surrogate/peer-rearing (SPR) and peer-rearing (PR) conditions increase the expression of genes involved in inflammation while decreasing expression of genes involved in type I IFN-mediated innate antiviral responses [i.e., the “conserved transcriptional response to adversity” previously observed in adults (9, 27–32)] and (ii) these effects are structured by transcription control pathways linked to stress-responsive “social signal transduction” pathways such as the sympathetic nervous system (SNS) and hypothalamus–pituitary–adrenal (HPA) axis (10, 13, 17, 32). These hypotheses are tested by coupling microarray-based assessment of the entire macaque transcriptome with recent advances in computational bioinformatics (33) and multiple-hypothesis testing (34–36) to map large ensembles of differentially-expressed genes into a small number of higher-order biological themes regarding their regulatory causes (e.g., transcription factor activity) (37), cellular contexts (e.g., originating leukocyte subtype) (30), and functional consequences (e.g., Gene Ontology functional annotations) (38).

Results

Effects of Surrogate/Peer Rearing. Previous studies have identified substantial increases in adult health risk in macaques exposed to SPR conditions in early life (5). Our initial analyses compare leukocyte transcriptional profiles in peripheral blood mononuclear cells (PBMCs) from 4 mo-old SPR animals ($n = 4$) relative to those from maternally reared (MR) animals ($n = 5$). Genome-

Author contributions: S.W.C., G.C., and J.J.H. designed research; S.W.C., G.C., J.M.G.A., A.M.R., and S.J.S. performed research; S.W.C., G.C., and J.J.H. contributed new reagents/analytic tools; S.W.C. and G.C. analyzed data; A.M.R. assisted with data collection; S.J.S. provided subjects; and S.W.C., G.C., J.J.H., and S.J.S. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE35850).

¹To whom correspondence may be addressed. E-mail: coles@ucla.edu or jjh@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1218253109/-DCSupplemental.

wide transcriptional profiling identified 249 transcripts showing a twofold or greater difference in average expression levels (85 up-regulated in SPR vs. MR and 164 down-regulated). Fig. 1 displays these transcriptome differences. **Dataset S1** lists specific up- and down-regulated genes.

Gene Ontology analyses (**Dataset S2**) characterize the genes up-regulated in PBMCs from SPR animals as being involved in multiple biological processes mediating immune activation, including metabolic activation (e.g., oxidation/reduction and glycogen metabolism), cytokine signaling (e.g., Interleukin 27, Leukemia Inhibitory Factor, Ciliary Neurotrophic Factor, Notch, and STAT1/STAT3 signaling), and T-cell proliferation. EntrezGene annotations for several up-regulated genes also indicate a common role in inflammation and tissue remodeling (e.g., **Dataset S1**: *IL8*, *MMP1*, *CCR3*, and *CCL2/MCP1*) (39). Down-regulated genes are characterized by involvement in innate immune response functions, such as antigen processing and presentation, antimicrobial defense responses (e.g., to bacteria and fungi), and liver development (**Dataset S2**). SPR down-regulated genes also include multiple transcripts identified in previous research as being involved in type I IFN-mediated innate antiviral responses (e.g., **Dataset S1**: *GBPI*, *IFIT1*, *IFIT2*, *IFITM3*, and *IRF7*) (39–42). Consistent with these functional bioinformatic results implicating T-cell activation and altered innate immune responses, transcript origin analyses identify monocytes and CD4⁺ T lymphocytes as major cellular sources of SPR up-regulated genes and B lymphocytes as major sources of SPR down-regulated genes (Fig. 1B).

To test the hypothesis that the observed differences in gene expression might be mediated by reciprocal up-regulation of proinflammatory NF- κ B family transcription factors and down-regulation of IFN response factor (IRF) family transcription factors, we carry out Transcriptional Element Listening System (TELiS) bioinformatic analysis of transcription factor-binding motifs in the promoters of differentially expressed genes. As in previous studies of social adversity (9, 27–29, 31, 32, 43), NF- κ B target motifs are significantly overrepresented within the promoter sequences of up-regulated genes and IRF motifs are significantly enriched within the promoter sequences of down-regulated genes (Fig. 1C). To determine whether such dynamics might be structured by social signal transduction pathways involving the SNS and the HPA axis (10, 13, 17, 32, 44), TELiS analyses also examine

motifs associated with the CREB transcription factors involved in β -adrenergic signaling by SNS catecholamines and glucocorticoid response elements (GREs) associated with HPA-axis signaling through the glucocorticoid receptor (GR). Results show significant enrichment of CREB sites within promoters of up-regulated genes, but no significant difference in GRE prevalence (Fig. 1C).

Effects of Peer Rearing. To determine whether similar dynamics might emerge in PR animals that lacked access to an inanimate surrogate mother and spent more time in contact with peers, we also compare PBMC gene expression profiles in PR animals ($n = 4$) vs. MR animals ($n = 5$). Analyses identify 256 transcripts showing a twofold or greater difference in average expression (105 up-regulated in PR vs. MR and 151 down-regulated; Fig. 2 and **Dataset S3**). Gene Ontology analyses again identify up-regulated genes as being involved in metabolic activation (ATP synthesis and electron transport chain, oxidation/reduction, and glycogen metabolism), gene translation, cytokine signaling (Interleukin 27, Leukemia Inhibitory Factor, Ciliary Neurotrophic Factor, Notch, and STAT1/STAT3 signaling), and T-cell proliferation (**Dataset S4**). Down-regulated genes are again characterized by involvement in innate immune response functions, including antigen processing and presentation and antimicrobial responses to bacteria, fungi, and viruses. Prominent among down-regulated transcripts are multiple genes involved in type I IFN-mediated innate antiviral responses (e.g., **Dataset S3**: *GBPI*, *IFI27*, *IFIT1*, *IFIT2*, *IFITM3*, *IRF7*, *MX1*, and *MX2*) (39–42). Gene Ontology annotations also link down-regulated genes to liver development and fat cell differentiation (**Dataset S4**). Transcript origin analyses identify monocytes and CD4⁺ T lymphocytes as cellular mediators of PR up-regulated genes (Fig. 2B) and B lymphocytes as cellular mediators of down-regulated genes. TELiS promoter-based bioinformatics again implicate a reciprocal increase in activity of NF- κ B and decrease in activity of IRF transcription factors in structuring the observed differences in immune response gene expression (Fig. 2C). TELiS analyses also indicate increased CREB activity, but provide no evidence of decreased GR-mediated transcription (Fig. 2C).

Comparison of Peer and Surrogate/Peer Rearing. The qualitative similarity in effects of SPR and PR conditions is underscored by the fact that relatively few genes show a twofold or greater

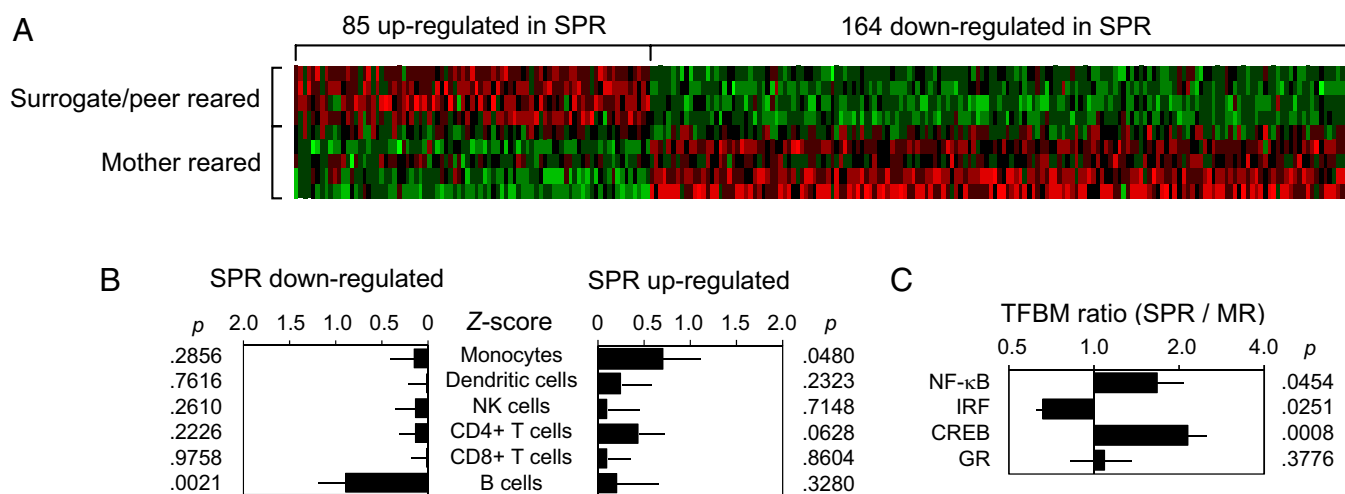


Fig. 1. Differential gene expression in leukocytes from mother-reared vs. surrogate/peer-reared macaques. (A) Heat plot representation of gene expression values for 249 transcripts showing a twofold or greater difference in average expression between peripheral blood mononuclear cells (PBMCs) from surrogate/peer-reared (SPR) rhesus macaques and those from maternally reared (MR) rhesus macaques. Rows, animals; columns, gene transcripts; red, up-regulated gene expression; green, down-regulated gene expression. (B) Transcript origin analyses assessing cellular origins of differentially expressed genes within specific PBMC subsets. (C) Relative prevalence of binding motifs for NF- κ B, IRF, CREB, and GR transcription factors within promoters of genes up-regulated in PBMCs from SPR vs. MR animals.

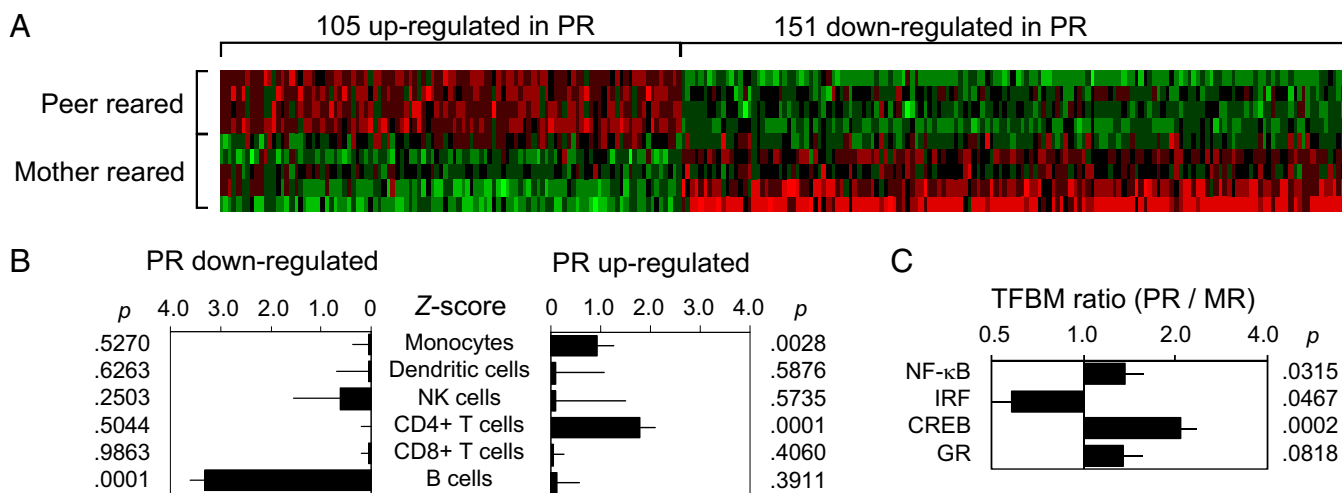


Fig. 2. Differential gene expression in leukocytes from mother-reared vs. peer-reared macaques. (A) Heat plot representation of gene expression values for 256 transcripts showing a twofold or greater difference in average expression between PBMCs from peer-reared (PR) rhesus macaques and those from maternally reared (MR) rhesus macaques. (B) Transcript origin analyses assessing cellular origins of differentially expressed genes within specific PBMC subsets. (C) Relative prevalence of binding motifs for NF- κ B, IRF, CREB, and GR transcription factors within promoters of genes up-regulated in PBMCs from PR vs. MR animals.

difference in average expression in direct comparison of these two groups ($n = 48$ up-regulated and 31 down-regulated, or less than one-third the number of differences observed in comparisons of each group with MR gene expression profiles; [Dataset S5](#)). Both SPR and PR groups show similar patterns of transcriptional differentiation from MR animals, with 49% (42/85) of SPR up-regulated transcripts also up-regulated by PR and 48% (79/164) of SPR down-regulated transcripts also down-regulated by PR (both exceeding the $<0.1\%$ overlap expected by chance; $P < 0.0001$). Gene Ontology analyses also identify few differences in the functional characteristics of SPR and PR PBMC transcriptomes ([Dataset S6](#)). Among the few functional differences that are identified is a comparative up-regulation of genes involved in antigen presentation (including proteolysis and antigen processing) in SPR animals relative to that in PR animals. No Gene Ontology annotations are identified as significantly up-regulated in PBMCs from PR animals relative to those from SPR animals.

Discussion

The results of this study show that adverse social conditions can become embedded in the basal transcriptome of primate immune cells within the first 4 mo of life. Compared with PBMCs from MR rhesus macaques, those from peer-reared animals (both SPR and PR) show enhanced expression of genes involved in inflammation and T-lymphocyte activation and reduced expression of genes involved in type I IFN-mediated innate antiviral responses and other pathogen-specific innate antimicrobial responses. This pattern of enhanced inflammatory gene expression and inhibited antiviral gene expression parallels the conserved transcriptional response to adversity (CTRA) observed in previous correlational studies of humans confronting adverse life circumstances (9, 27–32, 43). The experimental manipulation of early life social conditions in this study demonstrates that social adversity can play a causal role in activating CTRA dynamics and can do so during the earliest stages of postnatal immune system development. To the extent that such environmentally mediated transcriptome remodeling persists to affect immune responses to pathogens encountered later in life [e.g., inhibiting immune responses to viral infections (45, 46) or amplifying allergic inflammation (43)], the present findings provide a molecular framework for understanding the long-observed epidemiologic association between social adversity and reduced host resistance to disease (47–50), as well as more recently recognized effects of

early life social conditions on adult immune function (10, 12, 13, 51) and disease risk (1–5).

These data provide additional insights into the specific immune cell subtypes that are most sensitive to socio-environmental regulation and the neural and endocrine pathways that may mediate such relationships. Transcript origin analyses link SPR/PR-induced transcriptional up-regulation to monocytes and CD4⁺ T lymphocytes and transcriptional down-regulation to B lymphocytes. These findings parallel previous primate studies documenting altered CD4/CD8 T-lymphocyte ratios as a function of social vs. nonsocial housing conditions (52, 53) and defining leukocyte subset alterations as a key mechanism of social influences on the aggregate leukocyte transcriptome (54). These findings are also consistent with previous studies indicating monocyte-derived gene activation in humans confronting adversity (30, 31). On the basis of the known functions of these specific cell subtypes (30), SPR/PR animals might be expected to show reduced antibody responses (mediated by B cells, e.g., in response to vaccines or infections) and increased chronic inflammation (initiated by monocytes and perpetuated by CD4⁺ T lymphocytes, e.g., in responses to injury or infection). Additional research will be required to directly assess these specific immune system functional alterations, but the present results are broadly consistent with the increased disease risk observed in SPR/PR animals (5).

Also consistent with previous observations are results from promoter-based bioinformatic analyses implicating increased activity of proinflammatory NF- κ B transcription factors and decreased activity of IRF family transcription factors in structuring the observed gene expression differences (32). These analyses also implicate CREB family transcription factors as potential molecular mediators of PR/SPR effects on the basal leukocyte transcriptome. CREB factors play a central role in mediating the transcriptional effects of SNS activation via β -adrenergic receptors (55), and β -adrenergic signaling can also activate NF- κ B (56), up-regulate transcription of proinflammatory cytokine genes (57), and inhibit IRF transcription factors and type I IFN gene expression (46, 58) [i.e., the same proinflammatory/anti-antiviral transcriptional shift observed here and in other adversity studies (32)]. A potential increase in SNS-induced β -adrenergic signaling would parsimoniously account for many of the transcriptional dynamics observed here and provide a specific social signal transduction pathway by which early life social adversity alters basal leukocyte gene expression profiles and immune cell function (51, 52, 59, 60).

Parallel bioinformatic analyses provide little support for the hypothesis that reduced glucocorticoid-mediated transcription plays a major role in structuring the effects observed here [as has previously been observed in studies of long-term adversity in adults (9, 27, 28, 54)]. It is possible that the effects of early life adversity do not involve changes in cortisol signaling or GR function [e.g., as previously observed (12)]. However, it is also possible that GR-related functional alterations may have been present, but remain undetected due to the limited statistical power available in this study. Alternatively, the SNS/ β -adrenergic/CREB-related transcriptional dynamics observed here at 4 mo of age may constitute the initial phase of a long-term regulatory trajectory that subsequently induces GR desensitization and propagates systemic inflammation into adulthood. Ongoing longitudinal analyses of gene expression dynamics in the cohort studied in this paper will help clarify the possibilities. If analyses continue to indicate a role for SNS/ β -adrenergic/CREB signaling in structuring the CTRA gene expression dynamics observed here, pharmacologic inhibition of that pathway [e.g., with β -adrenergic antagonists (55–58)] might represent one potential strategy for mitigating the transcriptional dynamics and health risks associated with social adversity (13, 32).

Gene products that act recursively on the transcriptome (e.g., transcription factors and chromatin regulators) also show marked changes in expression in response to early life social adversity. Recursive dynamics can help to explain how transient periods of environmental adversity in early life propagate over time to influence adult health many years later. To the extent that transient environmental perturbations alter the expression of genes that control the basal dynamic equilibrium of the leukocyte transcriptome and/or alter the expression of molecules that mediate social signal transduction, early life social conditions may establish a long-lasting propensity to respond to challenges (either socio-environmental or microbial) that becomes manifest in health vulnerability only when the organism is challenged later in life (10–12). Quantitative modeling of such recursive developmental systems shows that corrective interventions are likely to be far more effective when deployed early in life than when delivered later in adulthood (16, 18–20, 61). The present study's observation that adverse social conditions can alter basal gene expression profiles in circulating immune cells within the first 4 mo implies an opportunity to remediate the biological impact of adverse environments shortly after they initially manifest and well before these gene regulatory regimes consolidate to drive the emergence of frank disease (e.g., late-life "diseases of aging" such as cardiovascular, neurodegenerative, or neoplastic disease) (62).

The scope of the present findings is limited in several important respects. First, this study does not include any direct measure of immune system functional activity [e.g., response to a pathogenic challenge (11, 12)], so the health significance of the present results remains to be determined in future studies. However, recent analyses have shown that macaques exposed to same adverse conditions in early life show significantly elevated physical and mental health problems in adulthood (5). The present study also focuses on circulating leukocytes, and implications for the more disease-relevant cells in peripheral tissues and lymphoid organs will require additional studies. Both of these limitations may be addressed to some extent by the fact that the type I IFN inhibition observed here parallels that observed in previous studies of social stress effects on macaque IFN responses to viral infection in lymph nodes (46). This study is also limited in its focus on a one-time analysis of immune cell gene expression after 4 mo of adverse social conditions in a small sample of animals. Replication of these findings in larger study samples with extended follow-up will be important to gauge the generalizability of these findings. Despite the fact that this small sample limits statistical power, the present analyses were nevertheless able to detect the proinflammatory and anti-antiviral transcriptional dynamics that have previously been observed in larger studies of social adversity in

adult humans and macaques (54). Future longitudinal analyses will be required to determine how rapidly such transcriptional alterations emerge, how long they persist, and whether or how quickly they might reverse following the cessation of environmental adversity. Previous analyses of the macaque SPR/PR paradigm have documented an enduring impact on other biobehavioral phenotypes (63), raising the possibility that the transcriptional alterations observed here may persist as well [e.g., due to self-perpetuating transcriptional dynamics (15–17)]. However, until more information is available regarding the longitudinal trajectory of the transcriptome dynamics documented here and their impact on cellular function and liability to disease, these findings should be considered the first steps toward addressing the question of molecular pathways by which early life social conditions affect subsequent health trajectories. Bioinformatic indications that SNS/ β -adrenergic/CREB activation may mediate the observed effects are consistent with previous experimental results (46, 56, 57), but future pharmacologic inhibition studies will also be required to decisively confirm the role of SNS signaling in the present paradigm. It is also important to note that the present findings can be interpreted only in the context of immune function, and the CNS neurobiological mechanisms of the present effects represent an important topic for future research (10, 21, 64, 65).

A major strength of this study is its ability to clearly map causal effects of early life social conditions on immune cell gene expression profiles in the context of a randomized experiment. Future analysis of larger samples and longitudinal trajectories from this paradigm hold great promise for clarifying the specific environmental conditions that mediate and moderate social genomic dynamics. For example, we observe particularly strong expression of IFN-related genes in MR animals reared in outdoor field cages (although laboratory-reared MR animals also showed elevations relative to SPR/PR animals). Such findings suggest that the physical environment may interact with the social environment to shape transcriptional adaptation. A larger sample of field-reared MR animals would be required for any definitive conclusion. The potential health significance of field vs. laboratory MR conditions also remains to be defined because previous epidemiologic analyses of this paradigm have excluded field-reared animals (5). Thus, ongoing longitudinal analyses of the rhesus macaque peer-rearing paradigm will provide valuable information with which to dissect the physical, behavioral, neural, and molecular pathways through which early life social conditions affect adult health and well-being (5).

Methods

Social Conditions. Newborn rhesus macaques were randomized to MR, PR, or SPR, as previously described in ref. 26. MR infants were housed in social groups approximating natural conditions (i.e., 8–10 adult females including the infant's mother, 2 adult males, and other similar-aged infants), with 2 of the 5 MR animals raised in 5-acre outdoor field cages and 3 reared in 14.6-m² indoor/outdoor laboratory enclosures (26). [Note that this distribution differs from the sample recently analyzed by Conti et al. (5), which did not include any MR animals reared in field cage environments.] Infants assigned to PR and SPR conditions were removed from their mothers between birth and 2 d of age and taken to a neonatal nursery, where they were housed individually with an inanimate surrogate mother for 37 d, after which they either entered a permanent group of four age-matched peers (PR) or continued individual housing with a surrogate mother supplemented by 2 h/weekday play sessions in groups of four age-matched peers (SPR) (26). (The PR monkeys also had a surrogate mother in the cage until they were 4 mo old.)

At 4 mo of age, peripheral blood mononuclear leukocytes (i.e., monocytes, dendritic cells, B lymphocytes, CD4⁺ T lymphocytes, CD8⁺ T lymphocytes, and NK cells) were isolated by standard Ficoll density gradient centrifugation of venipuncture blood samples obtained under resting conditions from a consecutive series of 14 healthy animals (8 females and 6 males) in the course of routine veterinary health examinations. Additional procedural details are provided in *SI Methods*. All procedures were carried out at the National

Institutes of Health Animal Center (Poolesville, MD) and approved by the Institutional Animal Care and Use Committee.

Gene Expression and Transcriptional Bioinformatics. Genome-wide transcriptional profiling and bioinformatic analyses were carried out essentially as described in previous studies (27, 30, 57), using Affymetrix Rhesus Genome Arrays. Detailed methods are provided in *SI Methods*, and data are deposited as Gene Expression Omnibus series GSE35850. Briefly, raw expression values for 52,024 probes assessing ~47,000 distinct macaque mRNA transcripts (including 7,185 distinct named macaque genes) were quantile normalized and \log_2 -transformed to identify genes showing twofold or greater differences in average expression levels across groups. Functional characteristics of differentially expressed genes were identified through National Center for Biotechnology Information (NCBI) EntrezGene annotations (66) and Gene Ontology (GO) annotations in the Affymetrix Rhesus Genome Array annotation file (i.e., both testing for significant differences in average levels of gene expression across rearing conditions, using small-sample permutation-based inference, and correcting for multiple-hypotheses testing using the stepdown algorithm of Romano and Wolf as developed and applied in ref. 34, and for overrepresentation of annotations in differentially expressed gene lists relative to the sampling frame of all genes present on the microarray, as outlined in *SI Methods*). Activity of specific transcriptional control pathways was assessed by TELIS bioinformatics analysis of transcription factor-binding motifs (TFBMs) in the promoters of differentially expressed genes (www.telis.ucla.edu) (37, 57). Promoter sequences derived from the *M. mulatta* genome

sequence were analyzed for TFBMs corresponding to NF- κ B (TRANSFAC V \$NFKAPPAB65_01 motif), IRFs (V\$IRF1_01), CREB (V\$CREB_01), and the glucocorticoid receptor (V\$GR_Q6), using PromoterScan and PromoterStats algorithms as previously described (37). Differential prevalence was quantified by the ratio of TFBM frequency in promoters of genes up-regulated in one group vs. another and summarized by the geometric mean ratio computed over nine parametric variations of promoter length (~300 bp relative to transcription start site, -600 bp, and -1,000 bp to +200 bp) and TFBM detection stringency (mat_sim = 0.85, 0.90, 0.95). Geometric mean ratios were tested for statistical significance, using a single-sample *t* statistic with bootstrap-derived SEs (37). Transcript origin analyses were conducted as previously described (30) to identify specific leukocyte subsets contributing to the observed PBMC transcriptome alterations.

ACKNOWLEDGMENTS. We thank Weihong Yan for parsing the *Macaca mulatta* genome to provide RefSeq promoter sequences, Kuan Chen and Yilan Chen for help in clustering the gene lists by GO biological function, and the National Institute of Aging (NIA) for catalyzing this research collaboration through the 2008 NIA Workshop on Genetic Methods and the Life Course. This research was supported by National Institutes of Health Grants P30AG107265, P60AG010415, R01HD054702, P30AG012857, R01AG034679, R01CA116778, and R37-HD065072; by funds from the Division of Intramural Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development; and by the American Bar Foundation. We acknowledge the support of a European Research Council grant hosted by University College Dublin, DEVHEALTH 269874.

- Kittleson MM, et al. (2006) Association of childhood socioeconomic status with subsequent coronary heart disease in physicians. *Arch Intern Med* 166(21):2356–2361.
- Galobardes B, Smith GD, Lynch JW (2006) Systematic review of the influence of childhood socioeconomic circumstances on risk for cardiovascular disease in adulthood. *Ann Epidemiol* 16(2):91–104.
- Cohen S, Doyle WJ, Turner RB, Alper CM, Skoner DP (2004) Childhood socioeconomic status and host resistance to infectious illness in adulthood. *Psychosom Med* 66(4):553–558.
- Gluckman PD, Hanson MA, Cooper C, Thornburg KL (2008) Effect of in utero and early-life conditions on adult health and disease. *N Engl J Med* 359(1):61–73.
- Conti G, et al. (2012) Primate evidence on the late health effects of early-life adversity. *Proc Natl Acad Sci USA* 109(23):8866–8871.
- Hertzman C (1999) The biological embedding of early experience and its effects on health in adulthood. *Ann N Y Acad Sci* 896:85–95.
- Levine S (2005) Developmental determinants of sensitivity and resistance to stress. *Psychoneuroendocrinology* 30(10):939–946.
- Zhang TY, et al. (2006) Maternal programming of defensive responses through sustained effects on gene expression. *Biol Psychol* 73(1):72–89.
- Miller GE, et al. (2009) Low early-life social class leaves a biological residue manifested by decreased glucocorticoid and increased proinflammatory signaling. *Proc Natl Acad Sci USA* 106(34):14716–14721.
- Lupien SJ, McEwen BS, Gunnar MR, Heim C (2009) Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat Rev Neurosci* 10(6):434–445.
- Avitsur R, Hunziker J, Sheridan JF (2006) Role of early stress in the individual differences in host response to viral infection. *Brain Behav Immun* 20(4):339–348.
- Nakamura T, et al. (2011) Maternal separation in early life impairs tumor immunity in adulthood in the F344 rat. *Stress* 14(3):335–343.
- Miller G, Chen E, Cole SW (2009) Health psychology: Developing biologically plausible models linking the social world and physical health. *Annu Rev Psychol* 60:501–524.
- Tyson JJ, Chen K, Novak B (2001) Network dynamics and cell physiology. *Nat Rev Mol Cell Biol* 2(12):908–916.
- Kim HD, Shay T, O'Shea EK, Regev A (2009) Transcriptional regulatory circuits: Predicting numbers from alphabets. *Science* 325(5939):429–432.
- Kauffman S (1993) *The Origins of Order: Self-Organization and Selection in Evolution* (Oxford Univ Press, Oxford).
- Cole SW (2009) Social regulation of human gene expression. *Curr Dir Psychol Sci* 18(3):132–137.
- Heckman JJ (2007) The economics, technology, and neuroscience of human capability formation. *Proc Natl Acad Sci USA* 104(33):13250–13255.
- Borghans L, Duckworth AL, Heckman JJ, ter Weel B (2008) The economics and psychology of personality traits. *J Hum Resour* 43(4):972–1059.
- Cunha F, Heckman JJ, Schennach S (2010) Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78(3):883–931.
- Weaver IC, Meaney MJ, Szyf M (2006) Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. *Proc Natl Acad Sci USA* 103(9):3480–3485.
- Champagne FA (2008) Epigenetic mechanisms and the transgenerational effects of maternal care. *Front Neuroendocrinol* 29(3):386–397.
- Borghol N, et al. (2012) Associations with early-life socio-economic position in adult DNA methylation. *Int J Epidemiol* 41(1):62–74.
- Essex MJ, et al. (2011) Epigenetic vestiges of early developmental adversity: Childhood stress exposure and DNA methylation in adolescence. *Child Dev*, 10.1111/j.1467-8624.2011.01641.x.
- Naumova OY, et al. (2012) Differential patterns of whole-genome DNA methylation in institutionalized children and children raised by their biological parents. *Dev Psychopathol* 24(1):143–155.
- Shannon C, et al. (2005) Maternal absence and stability of individual differences in CSF 5-HIAA concentrations in rhesus monkey infants. *Am J Psychiatry* 162(9):1658–1664.
- Cole SW, et al. (2007) Social regulation of gene expression in human leukocytes. *Genome Biol* 8(9, R189):R189.
- Miller GE, et al. (2008) A functional genomic fingerprint of chronic stress in humans: Blunted glucocorticoid and increased NF- κ B signaling. *Biol Psychiatry* 64(4):266–272.
- O'Donovan A, et al. (2011) Transcriptional control of monocyte gene expression in post-traumatic stress disorder. *Dis Markers* 30(2-3):123–132.
- Cole SW, Hawkley LC, Arevalo JM, Cacioppo JT (2011) Transcript origin analysis identifies antigen-presenting cells as primary targets of socially regulated gene expression in leukocytes. *Proc Natl Acad Sci USA* 108(7):3080–3085.
- Antoni MH, et al. (2012) Transcriptional modulation of human leukocytes by cognitive-behavioral stress management in women undergoing treatment for breast cancer. *Biol Psychiatry* 71(4):366–372.
- Irwin MR, Cole SW (2011) Reciprocal regulation of the neural and innate immune systems. *Nat Rev Immunol* 11(9):625–632.
- Cole SW (2010) Elevating the perspective on human stress genomics. *Psychoneuroendocrinology* 35(7):955–962.
- Heckman J, Moon SH, Pinto R, Savelyev P, Yavitz A (2010) Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics* 1(1):1–46.
- Lehmann EL, Romano JP (2005) *Testing Statistical Hypotheses* (Springer Science and Business Media, New York) 3rd Ed.
- Romano JP, Wolf M (2005) Exact and approximate stepdown methods for multiple hypothesis testing. *J Am Stat Assoc* 100(469):94–108.
- Cole SW, Yan W, Galic Z, Arevalo J, Zack JA (2005) Expression-based monitoring of transcription factor activity: The TELIS database. *Bioinformatics* 21(6):803–810.
- Beissbarth T, Speed TP (2004) GOSTat: Find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* 20(9):1464–1465.
- Jenner RG, Young RA (2005) Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol* 3(4):281–294.
- Ramilo O, et al. (2007) Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 109(5):2066–2077.
- Proud D, et al. (2008) Gene expression profiles during in vivo human rhinovirus infection: Insights into the host response. *Am J Respir Crit Care Med* 178(9):962–968.
- Waddell SJ, et al. (2010) Dissecting interferon-induced transcriptional programs in human peripheral blood cells. *PLoS ONE* 5(3):e9753.
- Chen E, et al. (2009) Genome-wide transcriptional profiling linked to social class in asthma. *Thorax* 64(1):38–43.
- McEwen BS (2007) Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiol Rev* 87(3):873–904.
- Shirtcliff EA, Coe CL, Pollak SD (2009) Early childhood stress is associated with elevated antibody levels to herpes simplex virus type 1. *Proc Natl Acad Sci USA* 106(8):2963–2967.
- Sloan EK, et al. (2007) Social stress enhances sympathetic innervation of primate lymph nodes: Mechanisms and implications for viral pathogenesis. *J Neurosci* 27(33):8857–8865.
- Virchow R (2006) Report on the typhus outbreak of Upper Silesia (1848). *Am J Public Health* 96(12):2102–2105.
- Cassel J (1976) The contribution of the social environment to host resistance: The Fourth Wade Hampton Frost Lecture. *Am J Epidemiol* 104(2):107–123.

Supporting Information

Cole et al. 10.1073/pnas.1218253109

SI Methods

Gene Expression and Transcriptional Bioinformatics. For genome-wide transcriptional profiling of circulating immune cells, 3-mL peripheral blood samples were obtained from a consecutive series of fourteen 4-mo-old infant rhesus macaques (mean age 4.5 mo, range 3.7–5.3, with at least 4 animals from each rearing condition: MR, PR, and SPR) as the animals were undergoing routine veterinary health examinations. Examinations included a period of chemical restraint induced by intramuscular injection of 10–20 mg/kg (0.1–0.2 mL/kg) ketamine hydrochloride, during which blood samples were collected from the femoral vein into EDTA-coated vacutainer tubes. Whole-blood samples were subject to standard Ficoll density gradient centrifugation (1) to isolate peripheral blood mononuclear cells (PBMCs; i.e., monocytes, dendritic cells, B lymphocytes, CD4⁺ T lymphocytes, CD8⁺ T lymphocytes, and NK cells). Total RNA was extracted from PBMC samples (RNeasy; Qiagen), tested for integrity by capillary electrophoresis (Agilent Bioanalyzer 2100), quantified by spectrophotometry (Nanodrop ND1000; Thermo Scientific), and assayed by Affymetrix Rhesus Genome Arrays according to the manufacturer's suggested protocol (Affymetrix), as previously described (2, 3). Thirteen of the 14 samples met basic quality criteria (RNA quantity and purity) and were retained for final analysis. Low-level expression measures for 52,024 probes assessing ~47,000 macaque gene transcripts (including 7,185 distinct named *Macaca mulatta* genes) were obtained from Affymetrix Expression Console software, quantile normalized using robust multiarray averaging (RMA) (4), and expressed as log₂-transformed values for statistical analysis. Differentially expressed genes were first identified by a twofold or greater difference in average expression value across groups, on the basis of previous results showing that fold-difference criteria yield more replicable gene lists than do *P* value-based criteria (5–10). Then, we used permutation-based inference procedures to address the problem of the small size of our sample, which called into question the validity of classical tests based on asymptotic inference. We further applied the general stepdown methodology of Romano and Wolf (11) as implemented in Heckman et al. (12) to assess family-wise error rates (FWER) defined over groups of differentially expressed genes showing common association according to specific Gene Ontology Biological Function annotations. Specifically, let the set of all relevant permutations of the treatment assignment *D* (i.e., MR, SPR, and PR) be represented by {*g*₁, . . . , *g*_{*J*}}. By relevant permutations we mean permutations that change the vector of treatment assignments; that is, *gD* ≠ *D*. Let *g*_{*J*+1} be the identity permutation. Permutation *P* values are based on the mid-*P* value of the prepivoting statistic, defined as

$$p \equiv \frac{1}{2(J+1)} \left(\sum_{l=1}^{J+1} \mathbf{1}[T^l \geq T^{J+1}] + \sum_{l=1}^{J+1} \mathbf{1}[T^l > T^{J+1}] \right),$$

where $T^l = \sum_{i=1}^{J+1} \mathbf{1}[\Delta^i \geq \Delta^l] / (J+1)$ is the prepivoted statistic and the Δ^l is the *t*-statistic computed for the permuted treatment assignment *g*_{*D*}. See Beran (13) for the statistical benefits of the prepivoting statistic. Under the randomization hypothesis, any real-valued statistic of the permuted data (i.e., *p*^{*j*}, *T*^{*j*}; *j* = 1, . . . , *J* + 1) that provides *J* + 1 distinct values as *g* varies in *G* is uniformly distributed across these *J* + 1 values. Lehmann and Romano (14) show that the mid-*P* value constitutes a valid level- α test; that is, Prob(Reject *H*₀) = Prob(*P* < α) = α . We use this test statistic in the stepdown algorithm of Romano and Wolf

(11) as implemented in Heckman et al. (12). Our type I error for the multiple-hypothesis testing is the FWER, which is the probability of rejecting any true null hypothesis. To illustrate the stepdown procedure, consider the null hypothesis of no treatment effect for a set of *K* joint outcomes, where the complement of this set is that there exists at least one hypothesis of *K* that we reject. The procedure starts by considering a joint test of all null hypotheses for the set of *K* hypotheses by comparing the maximum of the set of statistics associated with the hypotheses being jointly tested with the α -quantile of its distribution (α is the level of family-wise error rate that we want to control) to determine whether this first joint hypothesis is rejected or not. If we fail to reject the joint null hypothesis, then the algorithm stops; if we reject it, then we iterate and consider successive joint hypotheses that exclude the outcomes with the highest associated test statistics. Therefore, the procedure steps down, and at each successive step, it is implemented on a set of *K* – 1 null hypotheses. The process iterates until only one hypothesis remains. Romano and Wolf (11) show that the stepdown algorithm exhibits strong FWER control; that is, the FWER is held at or below a specified level regardless of the true configuration of the full set of hypotheses. Importantly, this procedure allows for the relaxation of the commonly made assumption in multiple-testing corrections of independence of the expression level of each gene of that of the other genes in the microarray.

Functional characteristics of differentially expressed genes were identified through National Center for Biotechnology Information (NCBI) EntrezGene RefSeq annotations (15) and differential representation analysis of Gene Ontology (GO) annotations (16) in the Affymetrix Rhesus Genome Array Annotation File. Differential representation analyses compared the proportion of differentially expressed genes tagged with a specific GO Biological Function annotation (i.e., $p_D = n_D/N_D$, with *n*_{*D*} representing the number of differentially expressed genes associated with a specific GO annotation, *N*_{*D*} representing the total number of differentially expressed genes, and *p*_{*D*} representing the proportion of differentially expressed genes tagged by that annotation) with the proportion observed across the sampling frame of all Rhesus Genome Array-assayed genes (i.e., $p_G = n_G/N_G$, with *n*_{*G*} representing the number of microarray-assayed genes associated with a specific GO annotation, *N*_{*G*} representing the 6,962 total assayed named and GO-annotated genes, and *p*_{*G*} representing the proportion of assayed genes associated with that annotation), using a one-tailed Fisher's exact test of equality of proportions.

To evaluate the role of specific transcription control pathways in structuring the observed pattern of differential gene expression, the promoters of differentially expressed genes were scanned by the Transcription Element Listing System (TELiS) transcription factor search engine (www.telis.ucla.edu) to identify differential representation of transcription factor-binding motifs (TFBMs) reflecting pathway activation (17). Promoter TFBMs were detected using TRANSFAC position-specific weight matrices (18) and the mat_sim match statistic (19). Promoter TFBM distributions were derived from the *M. mulatta* genome sequence, using PromoterScan and PromoterStats algorithms as previously described (3, 17). Differential TFBM prevalence was quantified by the ratio of frequencies in promoters of genes up-regulated in one experimental condition vs. another, averaged over nine different analytic stringencies involving parametric variations of promoter length (–300 bp relative to transcription start site, –600 bp, and –1,000 bp to +200 bp) and TFBM match

stringency (mat_sim = 0.85, 0.90, 0.95) (17). Average log ratios were tested for statistical significance, using a single-sample t test (after confirming that the distribution of log ratios was approximately Gaussian, as implied by the Central Limit Theorem) (20), and accuracy of the resulting P values was verified by a randomization test as previously described (2, 5). The randomization test involved 10,000 cycles in which (i) a random set of genes equivalent in number to those found to be empirically up-regulated and down-regulated in a given comparison was sampled from the population of macaque genes assayed by the Affymetrix Rhesus Genome Array and (ii) the TFBM prevalence ratio was calculated for each random sample. The two-tailed achieved significance level was computed as $2 \times (1 - p)$, where p denotes the percentile of the empirically observed log ratio within the null hypothesis distribution derived from 10,000 randomly sampled ratios. To identify specific leukocyte subsets contributing to the observed differences in PBMC gene expression, we conducted transcript origin analyses as previously described (21).

SI Datasets S1, S2, S3, S4, S5, and S6: Notes

Notes for Dataset S2: Gene Ontology Analysis and Single- and Multiple-Hypothesis Testing (Surrogate Peer-Reared vs. Mother-Reared Up-Regulated Genes). The first column in Dataset S2 identifies the cluster, defined according to the GO Biological Function of the gene. The second column provides the name of the probe set from the Rhesus Macaque Genome Array by Affymetrix. The third column provides the t -statistic based on the difference in means between the treatment and the control group. The fourth column shows the rank statistic associated with the t -statistic. We define this statistic below. Bigger values of the test statistic provide evidence against the null hypothesis H_0 of no treatment effects. The fifth column gives the one-sided single-hypothesis P value based on a small-sample permutation test. The last column provides the stepdown P value for the multiple-hypothesis testing according to ref. 11. Clusters of variables used in the multiple-hypothesis testing are separated by horizontal lines. We describe our inference method below.

Let the set of all relevant permutations of the treatment assignment D be represented by $\{g_1, \dots, g_J\}$. By relevant permutations we mean permutations that change the vector of treatment assignment; that is, $gD \neq D$. Let g_{J+1} be the identity permutation. Single P values in column 5 are based on the mid- P value of the prepivoting statistic, defined as

$$p \equiv \frac{1}{2(J+1)} \left(\sum_{l=1}^{J+1} \mathbf{1}[T^l \geq T^{J+1}] + \sum_{l=1}^{J+1} \mathbf{1}[T^l > T^{J+1}] \right),$$

where $T^l = \sum_{i=1}^{J+1} \mathbf{1}[\Delta^i \geq \Delta^l] / (J+1)$ is the prepivoted statistic and the Δ^j is the t -statistic computed for the permuted treatment assignment $g_j D$. See ref. 13 for the statistical benefits of the prepivoting statistic. Under the randomization hypothesis, any real-valued statistic of the permuted data (i.e., $p^j, T^j; j = 1, \dots, J+1$) that provides $J+1$ distinct values as g varies in G is uniformly distributed across these $J+1$ values. Ref. 14, Chap. 15, shows that the mid- P value constitutes a valid level- α test; that is, $\text{Prob}(\text{Reject } H_0) = \text{Prob}(P < \alpha) = \alpha$. We use this test statistic in the stepdown algorithm of ref. 11. Our type I error for the multiple-hypothesis testing is the FWER, which is the probability of rejecting any true null hypothesis. Ref. 11 shows that the stepdown algorithm has strong FWER control; that is, the FWER is held at or below a specified level regardless of the true configuration of the full set of hypotheses.

Notes for Dataset S2: Gene Ontology Overrepresentation Analysis. Surrogate peer-reared vs. mother-reared, GO annotations associated with up-regulated genes. The first column in Dataset S2 shows the GO

annotation for the biological function of the gene. The second column n_G shows the number of genes on the microarray with that function among those annotated with a biological process; the third column p_G shows their proportion, where $p_G = n_G/N_G$ and $N_G (= 6,962)$ is the total number of genes in the microarray annotated with a GO Biological Function category. The fourth column n_D shows the number of genes on the microarray with that function among those that are differentially expressed and have the GO Biological Function annotation; the fifth column p_D shows their proportion, where $p_D = n_D/N_D$ and $N_D (= 25)$ is the total number of up-regulated genes with a GO Biological Function annotation for the mother-reared vs. peer-reared comparison. The last column shows the one-sided P value for the Fisher's exact test for the equality of proportions against the alternative hypothesis of overrepresentation. Dataset S2 includes only annotations with $n_D > 1$.

Surrogate peer-reared vs. mother-reared, GO annotations associated with down-regulated genes. The first column in Dataset S2 shows the GO annotation for the biological function of the gene. The second column n_G shows the number of genes on the microarray with that function among those annotated with a biological process; the third column p_G shows their proportion, where $p_G = n_G/N_G$ and $N_G (= 6,962)$ is the total number of genes in the microarray annotated with a GO Biological Function category. The fourth column n_D shows the number of genes on the microarray with that function among those that are differentially expressed and have the GO Biological Function annotation; the fifth column p_D shows their proportion, where $p_D = n_D/N_D$ and $N_D (= 38)$ is the total number of down-regulated genes with a GO Biological Function annotation for the mother-reared vs. peer-reared comparison. The last column shows the one-sided P value for the Fisher's exact test for the equality of proportions against the alternative hypothesis of overrepresentation. Dataset S2 includes only annotations with $n_D > 1$.

Notes for Dataset S4: Gene Ontology Analysis and Single- and Multiple-Hypothesis Testing (Peer Reared vs. Mother Reared). The first column in Dataset S4 identifies the cluster, defined according to the GO Biological Function of the gene. The second column provides the name of the probe set from the Rhesus Macaque Genome Array by Affymetrix. The third column provides the t -statistic based on the difference in means between treatment and control group. The fourth column shows the rank statistic associated with the t -statistic. We define this statistic below. Bigger values of the test statistic provide evidence against the null hypothesis H_0 of no treatment effects. The fifth column gives the one-sided single-hypothesis P value based on a small-sample permutation test. The last column provides the stepdown P value for the multiple hypothesis testing according to ref. 11. Clusters of variables used in the multiple-hypothesis testing are separated by horizontal lines. We describe our inference method below.

Let the set of all relevant permutations of the treatment assignment D be represented by $\{g_1, \dots, g_J\}$. By relevant permutations we mean permutations that change the vector of treatment assignment; that is, $gD \neq D$. Let g_{J+1} be the identity permutation. Single P values in column 5 are based on the mid- P value of the prepivoting statistic, defined as

$$p \equiv \frac{1}{2(J+1)} \left(\sum_{l=1}^{J+1} \mathbf{1}[T^l \geq T^{J+1}] + \sum_{l=1}^{J+1} \mathbf{1}[T^l > T^{J+1}] \right),$$

where $T^l = \sum_{i=1}^{J+1} \mathbf{1}[\Delta^i \geq \Delta^l] / (J+1)$ is the prepivoted statistic and the Δ^j is the t -statistic computed for the permuted treatment assignment $g_j D$. See ref. 13 for the statistical benefits of the prepivoting statistic. Under the randomization hypothesis, any real-valued statistic of the permuted data (i.e., $p^j, T^j; j = 1, \dots, J+1$) that provides $J+1$ distinct values as g varies in G is

uniformly distributed across these $J + 1$ values. Ref. 14, Chap. 15, shows that the mid- P value constitutes a valid level- α test; that is, $\text{Prob}(\text{Reject } H_0) = \text{Prob}(P < \alpha) = \alpha$. We use this test statistic in the stepdown algorithm of ref. 11. Our type I error for the multiple-hypothesis testing is the FWER, which is the probability of rejecting any true null hypothesis. Ref. 11 shows that the stepdown algorithm has strong FWER control; that is, the FWER is held at or below a specified level regardless of the true configuration of the full set of hypotheses.

Notes for Dataset S4: Gene Ontology Overrepresentation Analysis. Peer-reared vs. mother-reared GO annotations associated with up-regulated genes. The first column in Dataset S4 shows the GO annotation for the biological function of the gene. The second column n_G shows the number of genes on the microarray with that function among those annotated with a biological process; the third column p_G shows their proportion, where $p_G = n_G/N_G$ and $N_G (=6,962)$ is the total number of genes in the microarray annotated with a GO biological function category. The fourth column n_D shows the number of genes on the microarray with that function among those which are differentially expressed and have the GO biological function annotation; the fifth column p_D shows their proportion, where $p_D = n_D/N_D$ and $N_D (=25)$ is the total number of up-regulated genes with a GO biological function annotation for the Mother-Reared vs. Peer-Reared Comparison. The last column shows the one-sided P value for the Fishers exact test for the equality of proportions against the alternative hypothesis of over-representation. The table only includes annotations with $n_D > 1$.

Peer-reared vs. mother-reared GO annotations associated with down-regulated genes. The first column in Dataset S4 shows the GO annotation for the biological function of the gene. The second column n_G shows the number of genes on the microarray with that function among those annotated with a biological process; the third column p_G shows their proportion, where $p_G = n_G/N_G$ and $N_G (=6,962)$ is the total number of genes in the microarray annotated with a GO biological function category. The fourth column n_D shows the number of genes on the microarray with that function among those that are differentially expressed and have the GO biological function annotation; the fifth column p_D shows their proportion, where $p_D = n_D/N_D$ and $N_D (=38)$ is the total number of down-regulated genes with a GO biological function annotation for the mother-reared vs. peer-reared comparison. The last column shows the one-sided P value for the Fisher's exact test for the equality of proportions against the alternative hypothesis of overrepresentation. Dataset S4 includes only annotations with $n_D > 1$.

Notes for Dataset S6: Gene Ontology Analysis and Single- and Multiple-Hypothesis Testing (Peer Reared vs. Surrogate Peer Reared). The first column in Dataset S6 identifies the cluster, defined according to the GO Biological Function of the gene. The second column provides the name of the probe set from the Rhesus Macaque Genome Array by Affymetrix. The third column provides the t -statistic based on the difference in means between the treatment and the control group. The fourth column shows the rank statistic

associated with the t -statistic. We define this statistic below. Bigger values of the test statistic provide evidence against the null hypothesis H_0 of no treatment effects. The fifth column gives the one-sided single-hypothesis P value based on a small-sample permutation test. The last column provides the stepdown P value for the multiple-hypothesis testing according to ref. 11. Clusters of variables used in the multiple-hypothesis testing are separated by horizontal lines. We describe our inference method below.

Let the set of all relevant permutations of the treatment assignment D be represented by $\{g_1, \dots, g_J\}$. By relevant permutations we mean permutations that change the vector of treatment assignment; that is, $gD \neq D$. Let g_{J+1} be the identity permutation. Single P values in column 5 are based on the mid- P value of the prepivoting statistic, defined as

$$P \equiv \frac{1}{2(J+1)} \left(\sum_{l=1}^{J+1} \mathbf{1}[T^l \geq T^{J+1}] + \sum_{l=1}^{J+1} \mathbf{1}[T^l > T^{J+1}] \right),$$

where $T^j = \sum_{l=1}^{J+1} \mathbf{1}[\Delta^l \geq \Delta^j] / (J+1)$ is the prepivoted statistic and the Δ^j is the t -statistic computed for the permuted treatment assignment $g_j D$. See ref. 13 for the statistical benefits of the prepivoting statistic. Under the randomization hypothesis, any real-valued statistic of the permuted data (i.e., $p^j, T^j; j = 1, \dots, J+1$) that provides $J+1$ distinct values as g varies in G is uniformly distributed across these $J+1$ values. Ref. 14, Chap. 15, shows that the mid- P value constitutes a valid level- α test; that is, $\text{Prob}(\text{Reject } H_0) = \text{Prob}(P < \alpha) = \alpha$. We use this test statistic in the stepdown algorithm of ref. 11. Our type I error for the multiple-hypothesis testing is the FWER, which is the probability of rejecting any true null hypothesis. Ref. 11 shows that the stepdown algorithm has strong FWER control; that is, FWER is held at or below a specified level regardless of the true configuration of the full set of hypotheses.

Notes for Dataset S6: Gene Ontology Overrepresentation Analysis (Surrogate Peer-Reared vs. Peer-Reared GO Annotations Associated with Up-Regulated Genes). The first column in Dataset S6 shows the GO annotation for the biological function of the gene. The second column n_G shows the number of genes on the microarray with that function among those annotated with a biological process; the third column p_G shows their proportion, where $p_G = n_G/N_G$ and $N_G (=6,962)$ is the total number of genes in the microarray annotated with a GO Biological Function category. The fourth column n_D shows the number of genes on the microarray with that function among those that are differentially expressed and have the GO Biological Function annotation; the fifth column p_D shows their proportion, where $p_D = n_D/N_D$ and $N_D (=25)$ is the total number of up-regulated genes with a GO Biological Function annotation for the mother-reared vs. peer-reared comparison. The last column shows the one-sided P value for the Fisher's exact test for the equality of proportions against the alternative hypothesis of overrepresentation. Dataset S6 includes only annotations with $n_D > 1$.

- Shannon C, et al. (2005) Maternal absence and stability of individual differences in CSF 5-HIAA concentrations in rhesus monkey infants. *Am J Psychiatry* 162(9):1658–1664.
- Conti G, et al. (2012) Primate evidence on the late health effects of early-life adversity. *Proc Natl Acad Sci USA* 109(23):8866–8871.
- Fuss IJ, Kanof ME, Smith PD, Zola H (2009) Isolation of whole mononuclear cells from peripheral blood and cord blood. *Curr Protoc Immunol* Chapter 7:Unit 7–1.
- Cole SW, et al. (2007) Social regulation of gene expression in human leukocytes. *Genome Biol* 8(9), R189/R189.
- Cole SW, et al. (2010) Computational identification of gene-social environment interaction at the human IL6 locus. *Proc Natl Acad Sci USA* 107(12):5681–5686.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19(2):185–193.

- Cole SW, Galic Z, Zack JA (2003) Controlling false-negative errors in microarray differential expression analysis: A PRIM approach. *Bioinformatics* 19(14):1808–1816.
- Shi L, et al.; MAQC Consortium (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24(9):1151–1161.
- Guo L, et al. (2006) Rat toxicogenomic study reveals analytical consistency across microarray platforms. *Nat Biotechnol* 24(9):1162–1169.
- Witten DM, Tibshirani R (2007) *A Comparison of Fold-Change and the t-Statistic for Microarray Data Analysis* (Stanford University, Stanford, CA), pp 1–17. Available at <http://www-stat.stanford.edu/~tibs/ftp/FCTComparison.pdf>.
- Shi L, et al. (2008) The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. *BMC Bioinformatics* 9(Suppl 9): S10.

