Supporting Information

Furman et al. 10.1073/pnas.1321060111

SI Materials and Methods

Serological Response to Trivalent Inactivated Seasonal Influenza Vaccine. The seroconversion rate (percent of individuals with a fourfold or greater change in their post- versus prevaccination microneutralization titer) was computed for each strain and for each group of individuals. The largest differences between males and females were observed for the H3N2 strain. Thus, the vaccine response was modeled as a binary variable (fold increase ≥4 to the H3N2 strain) in logistic regression analyses.

Gene Module Construction. First, gene probes were filtered by variance and normalized. Hierarchical agglomerative clustering was performed to derive 109 modules. For each gene module, a set of regulatory genes (regulatory program) was assigned based on regression analysis of genes in the modules onto expression of transcription factors. This was conducted using the LARS-EN algorithm (1). The LARS-EN algorithm provides fits of increasing numbers of predictors. To select the best model among the outputs of LARS-EN, we assessed the quality of the resulting models by the Akaike information criterion (AIC) (2), with sample-specific terms weighted by module variance. The fit with the best AIC score was selected for each module. Detailed statistical procedures have been described (3).

Interaction Analysis and Modeling of Vaccine Responsiveness. Potential confounders were identified if they modified the estimates of the sex effect on the vaccine response by more than 20%. A forward strategy was performed starting with a basic model including the sex covariate only.

To identify possible gene module candidates that explain the differences observed in vaccine response, we tested for marginal interactions between gene modules and the sex variable that associate with the neutralization antibody titer outcome (above). To do so, we used the Interact package (http://cran.r-project.org/web/packages/Interact/index.html), which searches for interactions in a binary response model using permutation methods to estimate false discovery rates (FDRs). The significance threshold was set at an FDR of <10% ($Q < 0.1$).

For the estimation of the regression coefficients and odds ratios in the response to vaccination, we conducted simple logistic regression with the categorical variable corresponding to the seroconversion to the H3N2 strain. The following formulas were used in the different models.

Model 1.

$$logit(y_i) = \mu + \beta_s male_i + \beta_c crp_i + \beta_{m42} mod\_42_i + \varepsilon_i$$

Model 2.

$$logit(y_i) = \mu + \beta_s male_i + \beta_c crp_i + \beta_{m42} mod\_42_i + \beta_{m52} mod\_52_i + \beta_{s:m52} male : mod\_52_i + \varepsilon_i$$

Model 3.

$$logit(y_i) = \mu + \beta_c crp_i + \beta_{m42} mod\_42_i + \beta_{m52} mod\_52_i + \beta_a age_i + \beta_{Tlo} maleTlo_i + \beta_{Thi} maleThi_i + \beta_{Tlo:m52} maleTlo : mod\_52_i + \beta_{Thi:m52} maleThi : mod\_52_i + \varepsilon_i,$$

where y_i is the binary response to H3N2 for the ith individual, μ is the average response for females, $male_i$ is a dichotomic variable (1 for male, 0 for female), crp_i is the CRP level of the ith individual, mod_42_i is the 42nd module median expression level of the ith individual, mod_52_i is the 52nd module expression level of the ith individual, $maleTlo_i$ is a dichotomic variable (1 for males with low levels of testosterone, that is, below the median of the male group; 0 for females and males with high levels of testosterone; see below), $maleTlo_i$:mod_52_i is the interaction term of mod_52_i and $maleTlo_i$ for the ith individual, $maleThi_i$ is a dichotomic variable (1 for males with high levels of testosterone, that is, above the median of the male group; 0 for females and for males with low levels of testosterone), $maleThi_i$:mod_52_i is the interaction term of mod_52_i and $maleThi_i$, and ε_i is the error term for the ith individual.

1. Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B* 67:301–320.
2. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19(6):716–723.
3. Furman D, et al. (2013) Apoptosis and other immune biomarkers predict influenza vaccine responsiveness. *Mol Syst Biol* 9:659.

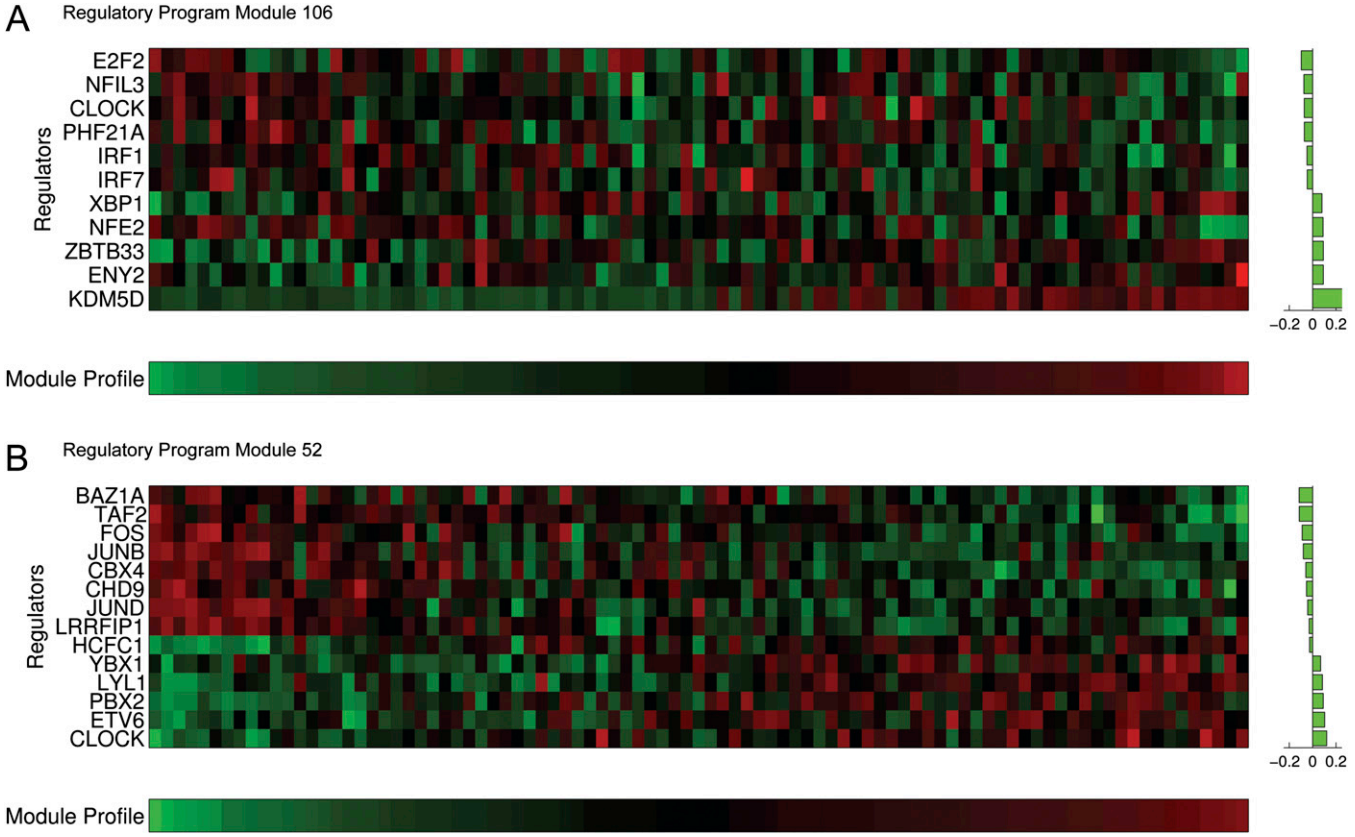

Fig. S1. Profile of gene regulators of modules 52 and 106. The initial gene expression data were reduced to gene modules by clustering analysis and assignment of a set of transcription factors (regulatory program) to each gene module. We used hierarchical agglomerative clustering to derive 109 modules. Using a set of candidate regulators composed of known signaling and transcription factors, for each gene module a set of regulatory genes (regulatory program) was assigned based on regression analysis of genes in the modules onto expression of transcription factors using the AIC (1). The regulatory program of module 52 contains FOS, JUNB, and JUD, among others, which is consistent with the suppressing effect of testosterone signaling on the AP-1 complex (FOS/JUN) (2). The regulatory program of sex-related gene module 106 contains transcription factors known to be differentially regulated in males versus females, such as CLOCK (3, 4), ENY2 (5), and IRF1 and IRF7 (6). Module profile, median expression of genes in the module.

1. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19(6):716–723.
2. Kallio PJ, Poukka H, Moilanen A, Jänne OA, Palvimo JJ (1995) Androgen receptor-mediated transcriptional regulation in the absence of direct interaction with a specific DNA element. *Mol Endocrinol* 9(8):1017–1028.
3. Gómez-Abellán P, et al. (2012) Sexual dimorphism in clock genes expression in human adipose tissue. *Obes Surg* 22(1):105–112.
4. Lim AS, et al. (2013) Sex difference in daily rhythms of clock gene expression in the aged human cerebral cortex. *J Biol Rhythms* 28(2):117–129.
5. Xiao R, et al. (2012) In utero exposure to second-hand smoke aggravates adult responses to irritants: Adult second-hand smoke. *Am J Respir Cell Mol Biol* 47(6):843–851.
6. Haslinger C, et al. (2004) Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. *J Clin Oncol* 22(19):3937–3949.

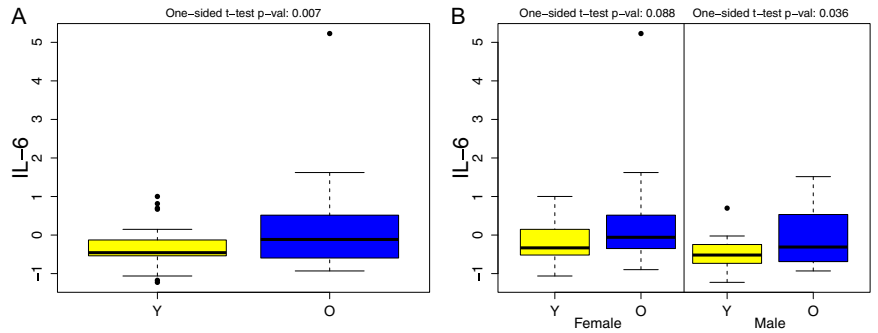

Fig. S2. Serum levels of IL-6 are increased in the elderly. To test for possibly explanatory variables of the elevated baseline levels of pSTAT3 proteins in blood monocytes in the elderly, we compared the serum IL-6 levels in all young (Y) versus older (O) individuals (*A*) or divided by sex (*B*). Significant differences are observed in all subjects (*A*), as well as in male individuals and to a lesser extent in females (*B*).

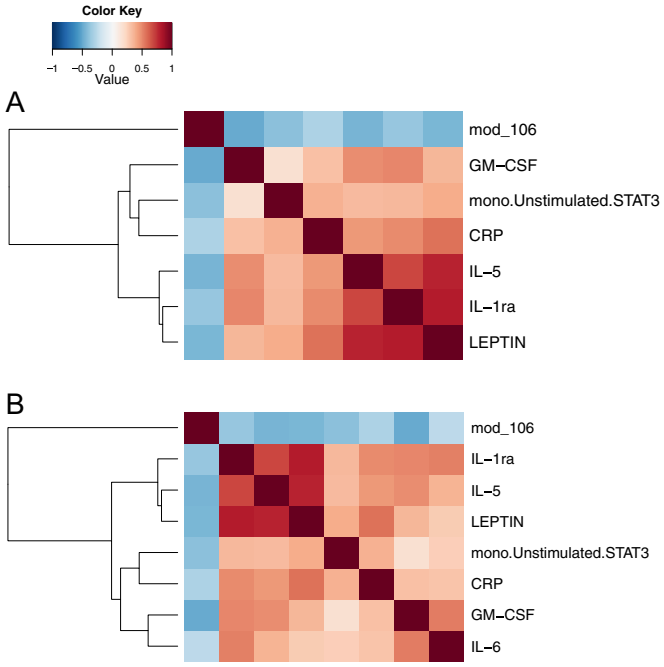

Fig. S3. Heat map of the correlation structure for sex-related immune features. A correlation matrix (Spearman method) was computed for all seven sex-related immune features in all individuals without (*A*) or with IL-6 (*B*), and hierarchical clustering (with Ward's method and Euclidian distance) was conducted. mono.Unstimulated.STAT3 clustered with CRP and to a lesser extent with GM-CSF (*A*, dendrogram), as well as with CRP and to a lesser extent with IL-6 and GM-CSF (*B*, dendrogram).

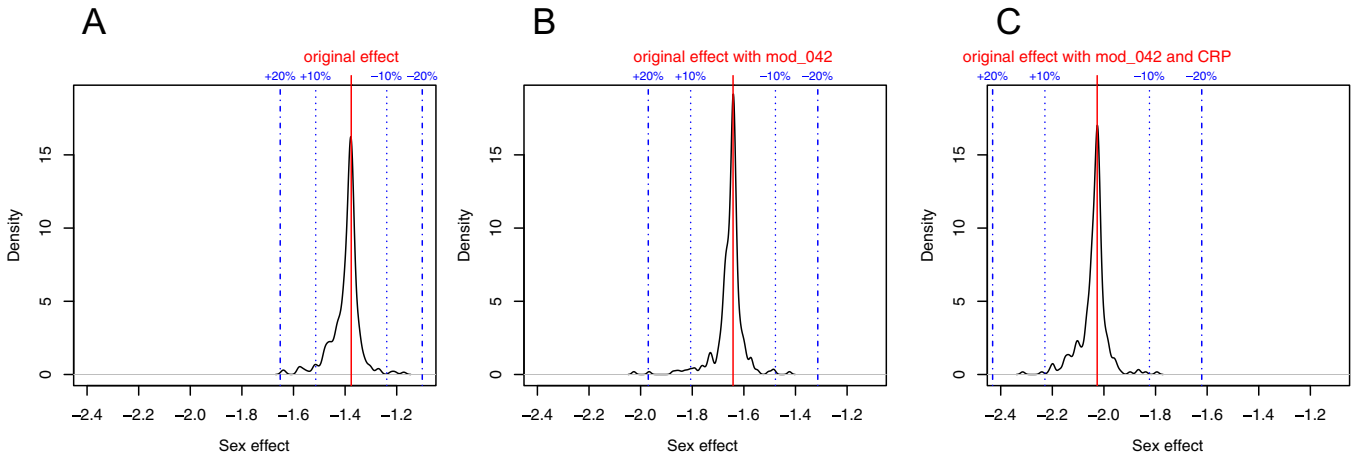

Fig. S4. Modifications in sex effect on vaccine response after adjusting for confounding factors. Forward stepwise logistic regression analysis was conducted to identify candidate confounders. (*A*) Regression coefficient of sex before adjustments. (*B*) Regression coefficient estimate for sex after adjusting for gene module 042. (*C*) Regression coefficient of sex after adjusting for gene module 042 and CRP levels. Negative values (*x* axis) indicate higher vaccine response in females.

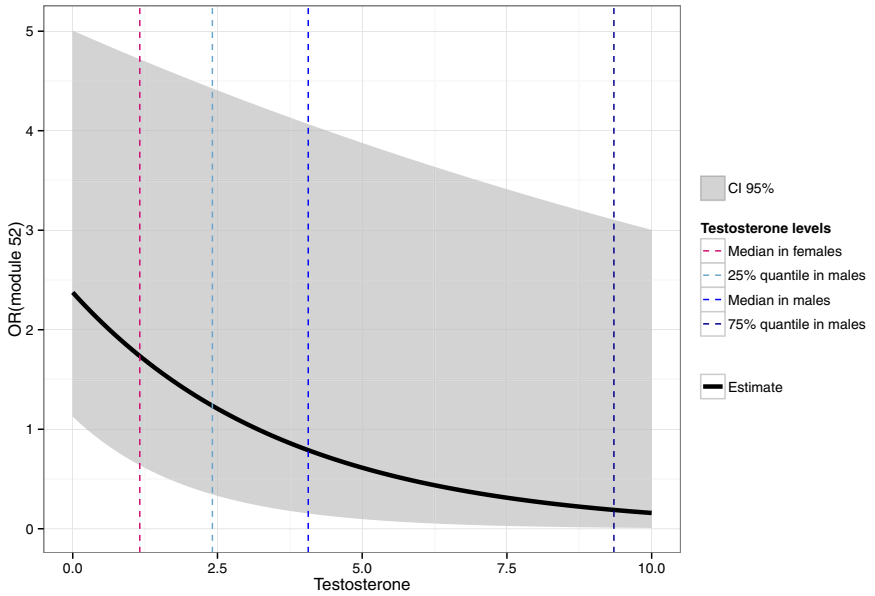

Fig. S5. Odds ratio for vaccine response based on expression of module 052 and testosterone levels. Logistic regression analysis was conducted on the antibody-neutralizing activity based on expression of genes in module 052 and the testosterone levels as a continuous measurement. The estimated odds ratio (OR) for the antibody-neutralizing response is shown (black continuous line). Red, light blue, blue, and dark blue dashed lines indicate the median testosterone levels in females and first, second, and third quartiles of testosterone levels for males. CI, confidence interval.

Table S1. Official gene symbol, Entrez ID, and module assignments for construction of gene modules 042, 052, and 106

Gene symbol	Entrez ID	Module assignment
RPS26P39	100128168	042
RPS26P38	100129552	042
RPS26P54	100131971	042
ZNF511	118472	042
SYT11	23208	042
GZMB	3002	042
RPS26P6	392256	042
RPS26P47	400156	042
ASCL2	430	042
RPS26P35	441377	042
RPS26P11	441502	042
ABI3	51225	042
CPSF3	51692	042
EXOSC10	5394	042
TRIT1	54802	042
PNPO	55163	042
RPS26	6231	042
RPS26P20	644166	042
RPS26P8	644191	042
RPS26P15	644928	042
RPS26P50	644934	042
RPS26P31	645979	042
RPS26P2	646753	042
RPS26P53	728823	042
RPS26P25	728937	042
CHRNA2	1135	052
FAM83F	113828	052
SPATA2L	124044	052
COX6C	1345	052
ZNF358	140467	052
CCDC140	151278	052
ADRA2C	152	052
AIM1	202	052
NAT9	26151	052
BSCL2	26580	052
GPR162	27239	052
C17orf60	284021	052
DHRS4L2	317749	052
HSPB1	3315	052
ANKRD33	341405	052
ARAF	369	052
FLJ41423	399886	052
RPS15P4	401019	052
LTA4H	4048	052
FAM116B	414918	052
MIF	4282	052
LOC440313	440313	052
LOC440993	440993	052
AURKAIP1	54998	052
USE1	55850	052
PDSS2	57107	052
PEX5	5830	052
RPS19	6223	052
BDKRB1	623	052
HSPBL2	653553	052
RPS19P3	728953	052
SPRYD3	84926	052
FIBCD1	84929	052
PIGQ	9091	052
NR1D1	9572	052
MTCP1NB	100272147	106
RPS4Y2	140032	106

Table S1. Cont.

Gene symbol	Entrez ID	Module assignment
CYorf15A	246126	106
NAAA	27163	106
NFU1	27247	106
GTF3A	2971	106
MTCP1	4515	106
PPA1	5464	106
PID1	55022	106
PLCXD1	55344	106
PRKY	5616	106
RPS4Y1	6192	106
CYorf15B	84663	106
ACCS	84680	106
DDX3Y	8653	106
EIF1AY	9086	106
KIAA0020	9933	106