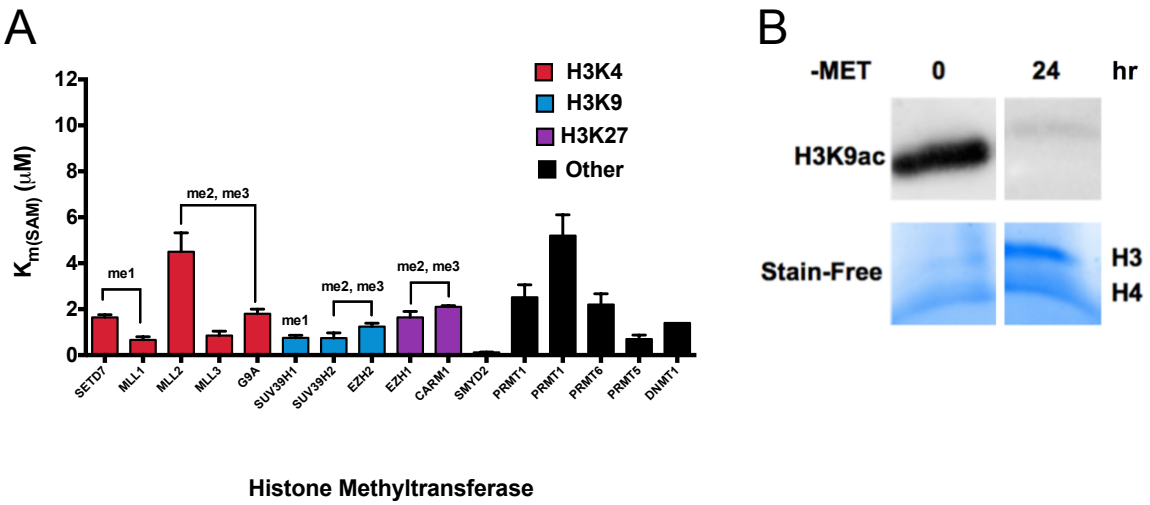
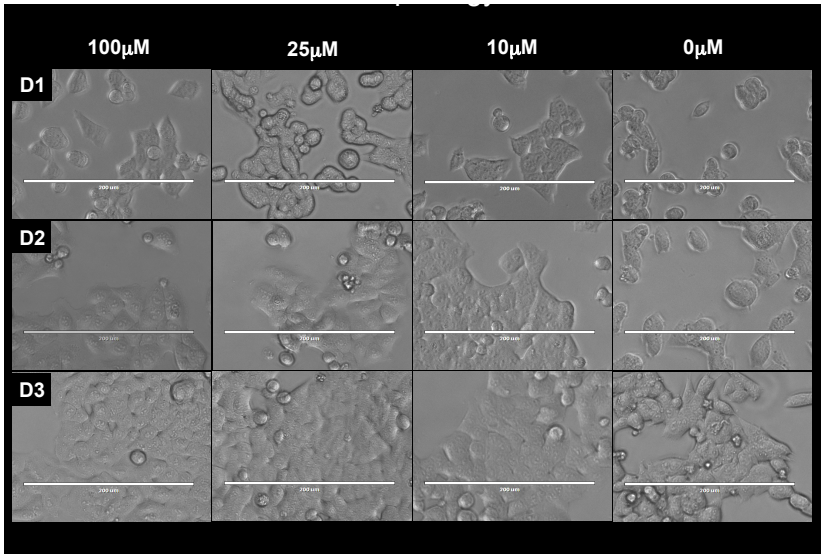
Figure S1, related to Figure 1.

Methionine metabolism alters histone methylation status
a.) Reported methyltransferase K_m values. b.) Changes in H3K9ac after methionine restriction c.) Cell morphology in response to methionine restriction at different days. e.) p53 response after methionine restriction or 5-FU treatment in Hct116.

Figure S1 (Related to Figure 1)





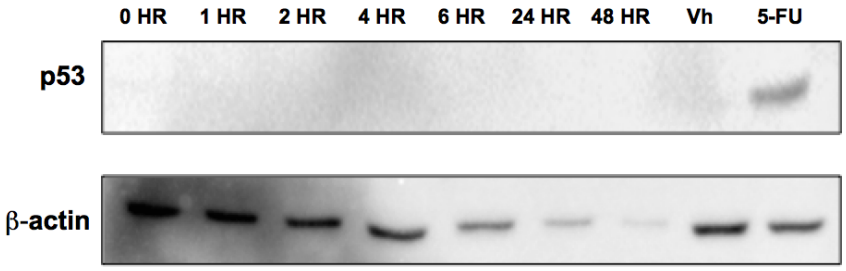


Figure S2, related to Figure 2

^{13}C Serine shows no contribution of the folate cycle to methionine metabolism
a.) ^{13}C serine tracing shows no contribution of folate metabolism to the methionine cycle. A map of serine carbon labels is shown. The levels of M+5 SAM and M+1 Methionine which denote flux into the methionine cycle that is derived from folate metabolism are negligible over natural abundance.

Figure S2 (related to Figure 2)

A

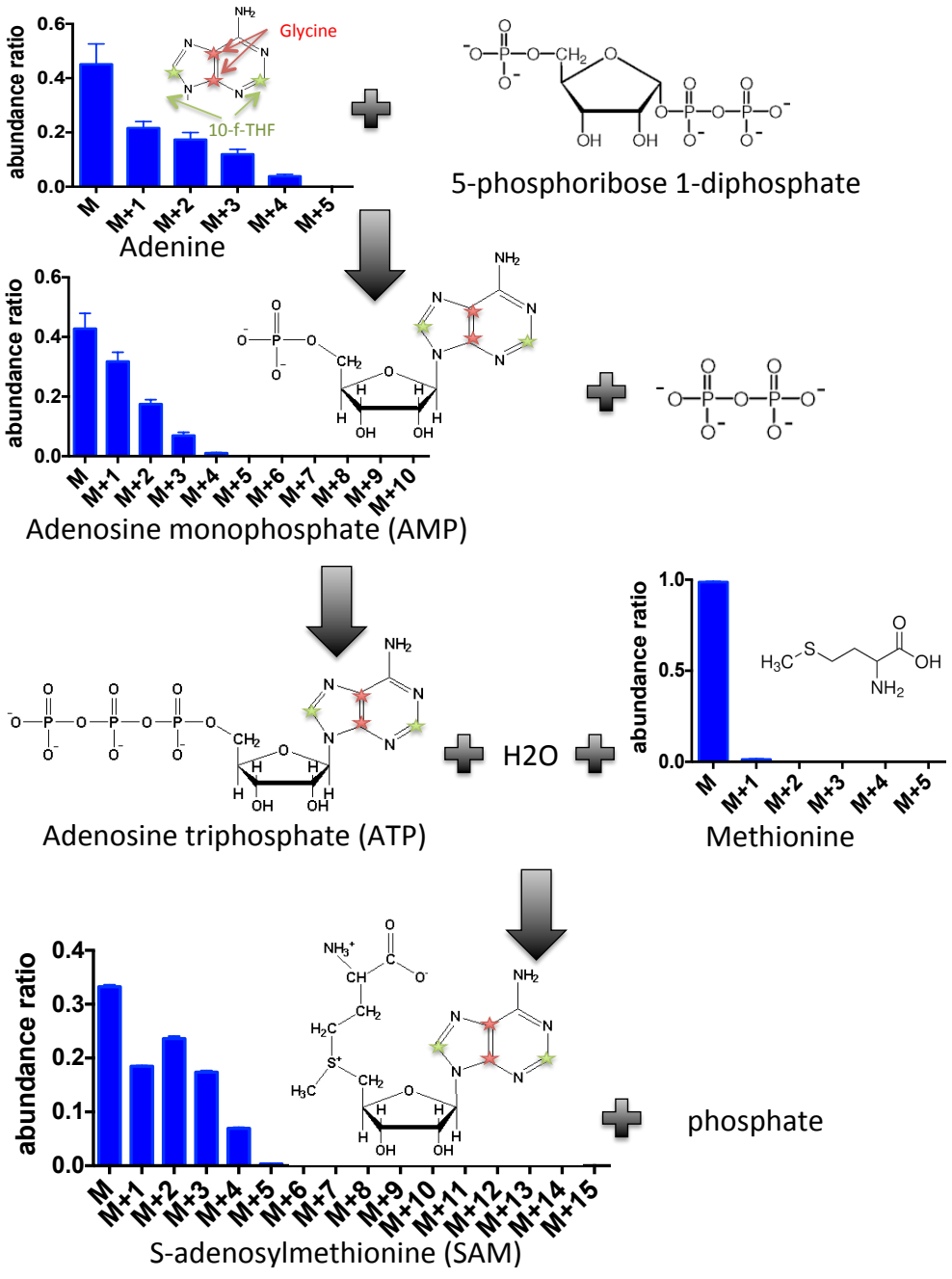


Figure S3, related to Figure 4

Methionine restrictions decreases H3K4me3 and alters gene expression
a.) Dot blot using H3K4me3 antibody with unmodified and modified histone peptides with varying concentrations. b.) H3K4me3 antibody specificity calculated from the histone peptide array. c.) Hierarchical clustering of genome-wide H3K4me3 peaks for replicate samples in methionine and methionine-restricted conditions. d.) Genome-wide binding site affinity from the TSS after methionine restriction e.) Differential (FDR < 0.1) H3K4me3 bound sites after methionine restriction. f.) Differentially expressed (FDR < 0.05) genes after methionine restriction. h.) Overlaid replicate H3K4me3 ChIP-seq signal with corresponding gene expression from RNA-seq. i.) Distribution of H3K4me3 at the TSS of colon cancer genes with significantly reduced expression after methionine restriction compared to the average genome distribution.

Figure S3 (related to Figure 4)

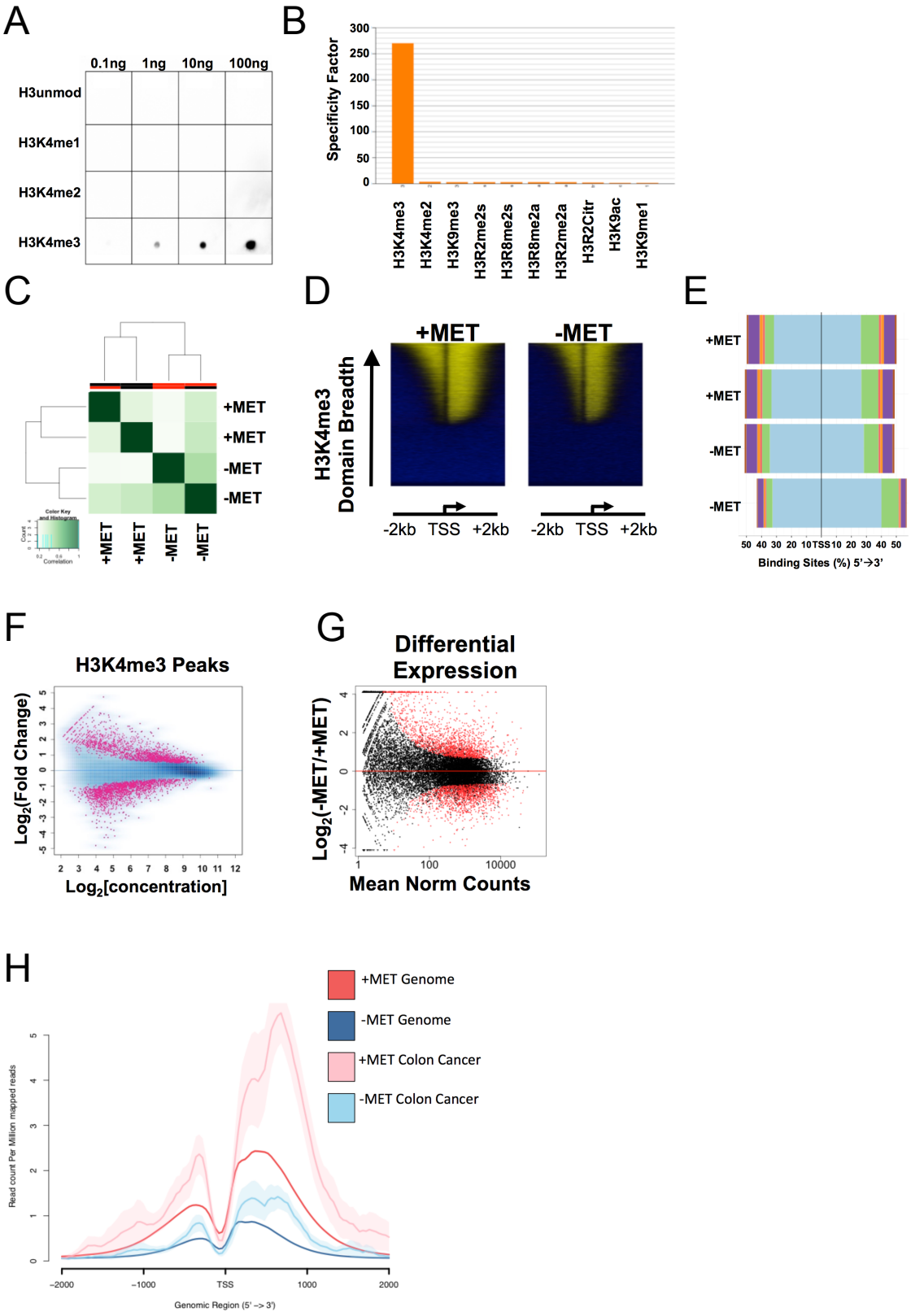


Figure S4, related to Figure 5

Alterations in methionine metabolism induce changes in histone methylation in vivo

a.) Global effects on metabolism in mouse serum and liver due to methionine restriction. b.) Pathway analysis on serum metabolites significantly different between control and MET restricted groups. c.) Pathway analysis on liver metabolites significantly different between control and MET restricted groups. d.) Quantile-quantile plot of differences in serum and liver. e.) Correlation analysis of metabolism in liver and plasma. f.) Correlation analysis of methionine levels and metabolites in plasma of MET-restricted animals. Significantly correlated metabolites are identified. All statistical analyses performed were student's t-tests, n=8, two-tailed, $\alpha<0.05$. Correlations of other one carbon metabolism related metabolites and H3K4me3 with liver (g) and plasma (h).

Figure S4 (related to Figure 5)

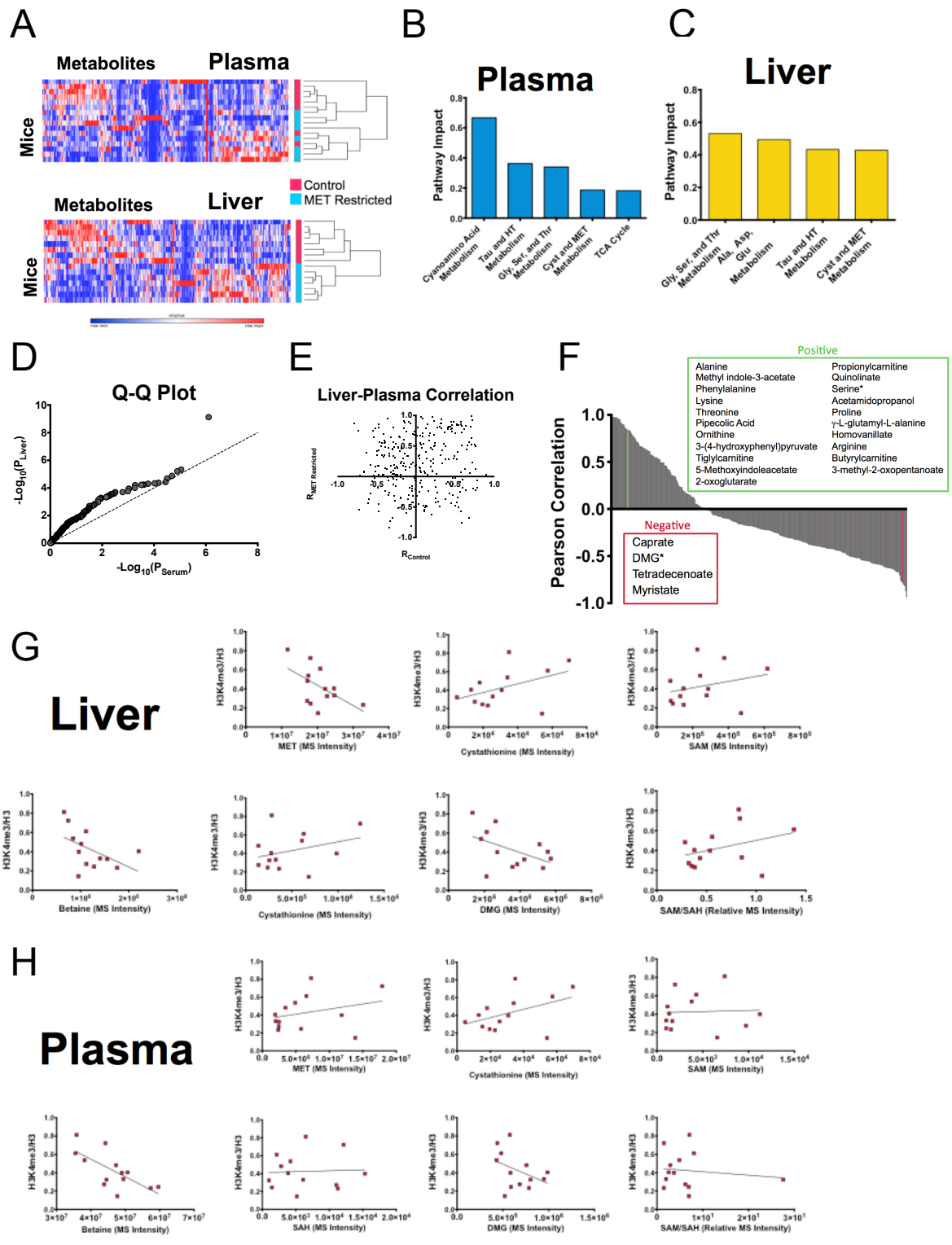


Figure S5, related to Figure 6

Methionine and metabolic variation in human subjects
a.) Measurement of a metabolomics profile across the human cohort (N=38). Results of unsupervised clustering of a distance matrix for metabolites are shown. b.) Pathways that correspond to clusters observed in (a). Pathways were identified from consideration of the highest pathway impacts of all metabolites contained in the specified cluster denoted by different colored boxes. c.) Spearman correlation coefficients of diet variables with methionine levels. d.) Spearman correlation coefficients of serum metabolites with methionine levels.

Figure S5 (related to Figure 6)

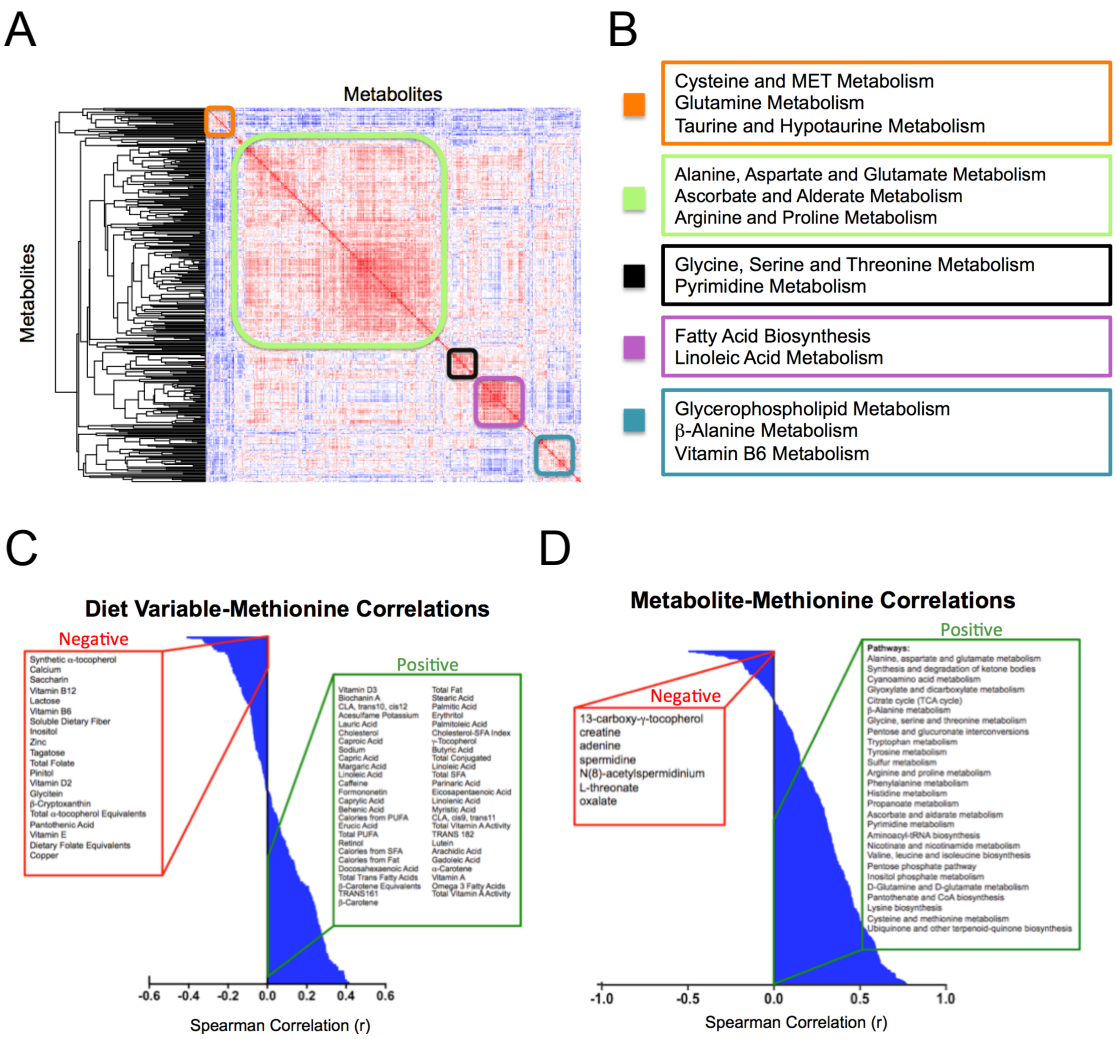
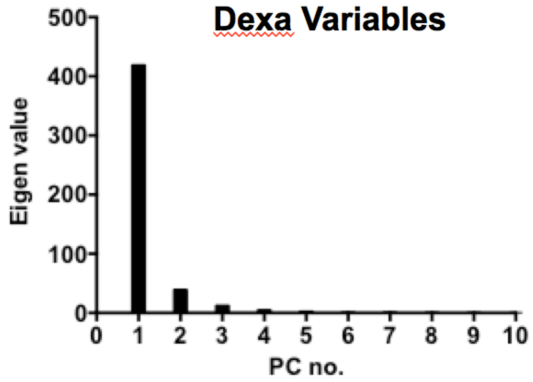


Figure S6, related to Figure 7

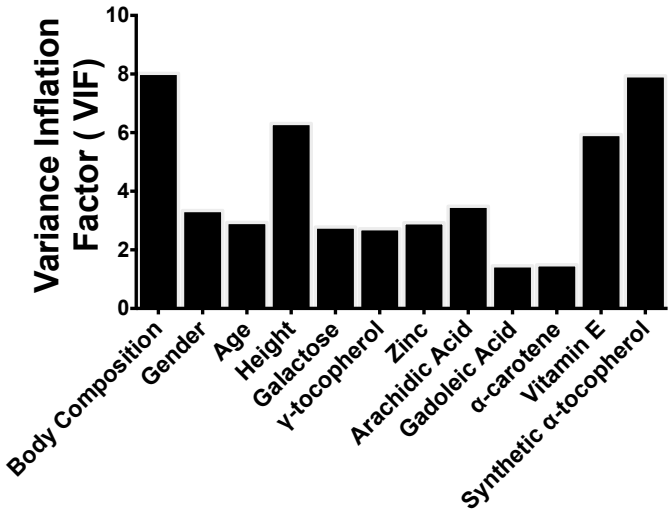
A computational model identifies determinants of methionine variability
a.) Eigenvalue spectrum form principal components analysis of DEXA variables. b.) Calculation of variance inflation factor for each variable.

Figure S6 (related to Figure 7)

A



B



Supplemental Experimental Procedures

Methionine Restriction in Cell Culture

For methionine deprivation, RPMI 1640 lacking amino acids, glucose, and glutamine was supplemented with Minimal Essential Media Non-Essential Amino Acids (MEM NEAA), 5mM glucose, 2 mM glutamine, 10% FBS, 100 U/mL penicillin, and 100 mg/mL streptomycin. Essential amino acids, except methionine, were added back at the same concentrations found in MEM Amino Acids Solution for minus methionine (-MET) media. For plus methionine (+MET) media, methionine was added to -MET media to a final concentration of 100μM. For titration experiments, methionine was added at the respective concentrations to the -MET media. Each additional amino acid dropout media was prepared in a similar manner but with concentrations supplemented to the respective condition in the plus amino acid condition. At the start of each experiment, cells were seeded at a density of 2.2×10^6 cells for 10cm plates for protein collection or 3×10^5 cells/well in a 6-well plate for metabolite collection and allowed to adhere for 24 hours. After a 24 hour incubation, cells were washed once with PBS and placed in respective media for 24 hours unless otherwise noted.

Mouse Feeding and Tissue and Plasma Analysis

Seven week-old C57BL/6J mice were purchased from Jackson Laboratories (Bar Harbor, ME) and housed in a conventional animal facility maintained at 20±2°C, 50±10% relative humidity and a 12 h light: 12 h dark light cycle. All animal procedures were approved by the Institutional Animal Care and Use Committee of the Orentreich Foundation for the Advancement of Science, Inc. (Permit Number: 0511MB). Food and water were provided *ad libitum*. Upon arrival, the mice were acclimated for one week and fed Rodent Chow 5001 (Purina) and water ad libitum. After one week, the mice were randomized and fed either an isocaloric 0.84% (w/w) methionine (control-fed; CF, N=8) diet or 0.12% (w/w) methionine (methionine-restricted; MR, N=8) diet from Research Diets (New Brunswick NJ). After 12 weeks of feeding, mice were fasted for 4 hours at the beginning of the light cycle to establish a physiological baseline and then sacrificed. Blood was collected from the retro-orbital plexus and plasma was collected, flash frozen and stored at -80°C until analyzed. The liver was harvested, flash frozen and stored at -80°C until processed. Liver tissues samples were fixed in 10% formalin solution (Thermo Scientific) and paraffin embedded. Hematoxylin and eosin (H & E) staining and imaging was carried out at Histowiz (New York, NY), high resolution slide scanning was carried out using Leica SCN400 and image analysis was conducted using Image-pro software (Media Cybernetics).

Immunoblotting

Samples were homogenized in Triton Extraction Buffer (TEB: 0.5% Triton X 100, 2mM phenylmethylsulfonyl fluoride (PMSF), 0.02% NaN_3 in PBS) and centrifuged at 2000rpm for 10 min at 4°C. Pellets were resuspended in 0.2N HCl and histones were acid extracted overnight at 4°C. Histone extracts were loaded onto 12% SDS-PAGE gels and transferred to polyvinylidene difluoride (PVDF) membranes. Membranes were blocked in 5% dry milk in TBST and incubated with anti-H3K4me3 (Millipore) 1:2000, anti-H3K9me3 (Millipore) 1:2000, anti-H3K27me3 (Millipore) 1:10000, or anti-H3 (Millipore) 1:10000. Horseradish-peroxidase-conjugated anti-rabbit (Rockland) or anti-mouse (Rockland), 1:10,000, were used as secondary antibodies. Chemiluminescent signals were detected with Clarity Western ECL Detection Kit (Bio-Rad) and imaged using a ChemiDoc MP System (Bio-Rad). All western blots are normalized to the total H3 signal before relative quantification was calculated.

Metabolite Extraction

For culture from adherent cells, the media was quickly aspirated and cells were washed with cold PBS on dry ice. Then, 1mL of extraction solvent (80% methanol/water) cooled to -80°C was added immediately to each well, and the dishes were transferred to -80°C for 15 min. Plates were removed and cells were scraped into the extraction solvent on dry ice. For tissue, the sample was homogenized in liquid nitrogen and then 5 to 10 mg was weighed in a new Eppendorf tube. Ice-cold extraction solvent (250µL) was added to each tissue sample and homogenized using a tissue homogenizer. The homogenate was incubated on ice for 10 min. For plasma or serum, 20µL was transferred to a new Eppendorf tube containing 80µL HPLC grade water. Next, 400µL of ice-cold methanol was added to the sample for a final methanol concentration of 80% (v/v). Samples were incubated on ice for 10 min. All metabolite extracts were centrifuged at 20,000g at 4°C for 10 min. Finally the solvent in each sample was evaporated in a Speed Vacuum for metabolomics analysis. For polar metabolite analysis, the cell extract was dissolved in 15µL water and 15µL methanol/acetonitrile (1:1 v/v) (LC-MS optima grade, Thermo Scientific). Samples were centrifuged at 20,000g for 10min at 4°C and the supernatants were transferred to Liquid Chromatography (LC) vials. The injection volume for polar metabolite analysis was 5µL.

Liquid Chromatography

An Xbridge amide column (100 x 2.1 mm i.d., 3.5 µm; Waters) is employed on a Dionex (Ultimate 3000 UHPLC) for compound separation at room temperature. The mobile phase A is 20 mM ammonium acetate and 15 mM ammonium hydroxide in water with 3% acetonitrile, pH 9.0 and mobile phase B is acetonitrile. Linear

gradient as follows: 0 min, 85% B; 1.5 min, 85% B, 5.5 min, 35% B; 10min, 35% B, 10.5 min, 35% B, 14.5 min, 35% B, 15 min, 85% B, and 20 min, 85% B. The flow rate was 0.15 ml/min from 0 to 10 min and 15 to 20 min, and 0.3 ml/min from 10.5 to 14.5 min. All solvents are LC-MS grade and purchased from Fisher Scientific.

Mass Spectrometry

The Q Exactive MS (Thermo Scientific) is equipped with a heated electrospray ionization probe (HESI), and the relevant parameters are as listed: evaporation temperature, 120 °C; sheath gas, 30; auxiliary gas, 10; sweep gas, 3; spray voltage, 3.6 kV for positive mode and 2.5 kV for negative mode. Capillary temperature was set at 320°C, and S-lens was 55. A full scan range from 60 to 900 (m/z) was used. The resolution was set at 70 000. The maximum injection time was 200 ms. Automated gain control (AGC) was targeted at 3,000,000 ions.

Metabolomics and Data Analysis

Raw data collected from LC-Q Exactive MS is processed on Sieve 2.0 (Thermo Scientific). Peak alignment and detection are performed according to the protocol described by Thermo Scientific. For a targeted metabolite analysis, the method "peak alignment and frame extraction" is applied. An input file of theoretical m/z and detected retention time of 263 known metabolites is used for targeted metabolites analysis with data collected in positive mode, while a separate input file of 197 metabolites is used for negative mode. M/Z width is set at 10 ppm. The output file including detected m/z and relative intensity in different samples is obtained after data processing. Dot plots and other quantitation and statistics were calculated and visualized with the Graphpad prism software package.

RNA Sequencing

Total RNA was extracted using the PARIS kit (Life Technologies, cat#AM1921), polyA selected before library generation, and prepared according to a standard protocol provided by Illumina by the Weill Cornell Medical College Epigenomics Core. Samples were pooled into one lane on the Illumina HiSEQ2500 and sequenced. Reads were mapped to the reference genome (hg19) using TopHat2.0.7 in the Galaxy software suite (https://usegalaxy.org). For the analysis, we generated a counts table from accepted hits bam files using HTseq (Anders et al., 2015) with the following parameters: htseq-count –f bam –s yes [input bam] [hg19genes.gtf]. Differential expression was calculated from the counts table using the edgeR package in R (Nikolayeva and Robinson, 2014; Robinson et al., 2010) and DEseq2 (Love et al., 2014).

Chromatin Immunoprecipation and Sequencing

Hct116 cells were cultured in 15cm plates resulting in ~$3x10^7$ cells at 80% confluency. Cells were cross-linked for 5 minutes at room temperature by addition

of 1% formaldehyde to the cell culture media. The reaction was quenched for 5 minutes by the addition of 2.5M glycine. Cells were collected and lysed in FA buffer (50mM HEPES/KOH pH 7.5, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1M NaCl) with protease inhibitors. Cross-linked chromatin was resuspended in LB3 (10mM Tris-HCl, pH 8.0, 200mM NaCl, 1mM EDTA, pH 8.0, 0.5mM EGTA, pH 8.0, 0.1% Sodium deoxycholate, 0.5% N-lauroylsarcosine) and samples were sheared using the Covaris M2 ultrasonicator (Covaris INC, Woburn MA) . Each IP was performed as previously described, (Ercan et al., 2007; Landt et al., 2012) using $1.5x10^7$ cells and the Millipore ChIPAb+ H3K4me3 antibody (cat #17-614, lot #2196044) with Protein A Agarose beads (Millipore, cat #16-125, lot #2444123). Libraries were made and pooled according to Illumina instructions, then sequenced on an Illumina HiSEQ 2500 Rapid Run sequencer. Reads were mapped to the reference genome (hg19) using Bowtie2.2.4. H3K4me3 ChIP peaks were called using MACS2.08 with the broad setting. Alignment and peak calling were performed on the Galaxy platform (Blankenberg et al., 2010; Giardine et al., 2005; Goecks et al., 2010). H3K4me3 distribution was visualized using the ngsplot package (Shen et al., 2014) and ChIPseeker in R (Yu et al., 2015).

Analysis of Chip-seq Results for Differentially Bound Chromosome Regions

We used the R package "DiffBind" (Stark and Brown, 2011) to find regions with a significant difference in H3K4me3 between the minus-Methionine and the plus-Methionine conditions. In summary, peaks were first called by MACS2 to generate a consensus peak-set which we defined as a set of peaks called in at least 2 samples. Then using the original read files, we generated the counts table including the number of reads corresponding to each consensus peak in each sample. Finally using edgeR and DESeq2 we performed differential binding analysis on all of the consensus peaks and reported the significantly differentially bound regions (FDR<0.1). Finally, the "quantsmooth" package in R (Oosting et al., 2014; Ross-Innes et al., 2012) was used to overlay counts onto the human chromosomal ideograms.

Clinical Nutrition Studies

A four-day diet record was collected according to previous standards of recording diet to reflect habitual behaviors as has been previously described(Levine et al., 2014). Surveys were converted into nutritional variables according to previously described procedures using standard software(Levine et al., 2014). Nutrition data system for research (NDSR) dietary analysis software (University of Minnesota), a comprehensive food and nutrient database, was used to determine the average macro- and micronutrients consumed over the four-day period as previously described(Thalacker-Mercer et al., 2009). Before serum was collected, each subject was subjected to overnight fasting. Metabolites from serum were extracted using a previously described protocol(Shestov et al., 2014). All additional clinical variables were recorded according to previously described methods(Thalacker-Mercer et al., 2013; Thalacker-Mercer et al., 2009). The computational model of the factors contributing to methionine concentration variation is described in detail later.

Computational Modeling

Detailed Description

The dataset consists of 24 samples of serum methionine concentration (SMC) and 233 predictor variables. To understand what factors determine SMC, we develop a linear mixed model (a type of linear regression analysis) following the steps:

The final model:

$$y_i = \beta_0 + \sum_k \beta_k X_i^k + Z_i + \varepsilon_i,$$

where $i = 1, \dots, 24$ indexes the samples, y_i denotes SMC for the i-th sample, X_i^k ($k = 1, \dots, 24$) represent the 12 predictor variables that we model as fixed effect (intercept, alpha carotene, galactose, gamma tocopherol, gadoleic acid, arachidic acid, synthetic alpha tocopherol, vitamin E, zinc, age, gender, height, body composition), Z_i denotes the i-th sample that we model as random effect, and $\beta_0, \beta_1, \beta_2, \beta_3$ and ϵ_i are the standard parameter and noise terms in linear regression.

Variable Selection

Since the number of predictors vastly outnumber the sample size and presumably only a small subset of the predictors are related to SMC, we first have a filtering step to pick out the predictors that significantly correlate with SMC. To do this, each of the 233 predictors is individually fit to the following linear mixed model. A p-value cutoff of 0.05 for the significance of correlation is used, after which 34 variables are left; no correction for multiple testing is conducted in order to include as many potentially relevant variables as possible. We mention that a linear mixed model, rather than an ordinary linear model, is chosen to account for the fact that the some samples are not independent but come from the same individuals.

21 among the 34 variables are Dual-energy X-ray absorptiometry (DEXA) variables that measure bone mineral density. An inspection of the correlation matrix and the principal components analysis (PCA) spectrum revealed that they are highly correlated with each other. Therefore, we condensed each of the 21 variables into a single one by using the first principle component. Large gaps between the leading and following eigenvalues in the PCA spectrum suggests that little information is lost in such a condensation.

Among the remaining 14 variables, we check if there is substantial collinearity present. Collinearity refers to the situation in which some predictors contain heavily redundant information and causes problems for parameter identifiability and model interpretation. To check collinearity, we regress each of the predictors on all other predictors: if any one predictor is regressed well on others, then it can be explained well by others and can hence be removed from the model. We find that

none of our predictors can be regressed on others with an R^2 greater than 0.9 (corresponding to an often suggested threshold of 10 for variance inflation factor), which means no substantial collinearity is present in our 14 variables and each of them captures largely independent information.

Finally we perform a standard procedure of variable selection to achieve a good balance of model predictive power and simplicity. We follow the suggestion of Faraway and use the Akaike information criterion (AIC) as our criterion for model selection: the model with the lowest AIC is selected. With only 14 candidate variables, we perform an exhaustive evaluation of the AIC of all possible models.

Model Analysis

For a multiple linear regression $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$, a common interpretation is of a geometric kind: given vector $\boldsymbol{y}$, find a linear combination of the column vectors of $\boldsymbol{X}$, $\hat{\boldsymbol{y}} = \boldsymbol{X\beta}$, such that $\|\boldsymbol{y} - \hat{\boldsymbol{y}}\|$ is the smallest, or equivalently, $\boldsymbol{y} - \hat{\boldsymbol{y}}$ is orthogonal to $\hat{\boldsymbol{y}}$. When the set of predictors has a natural partition, such as diet vs. non-diet variables in our case, the model can be written as $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon}$, where $\boldsymbol{X\beta}$ and $\boldsymbol{Z\gamma}$ correspond to the two groups of predictors.

The goodness-of-fit by the linear model is often measured by the so-called *coefficient of determination*, R^2, which also has a geometric interpretation: let $SST = \sum_i (y_i - \overline{\boldsymbol{y}})^2 = \|\boldsymbol{y} - \overline{\boldsymbol{y}}\|^2$ be the total sum of squares (variation), it can be shown that $\|\boldsymbol{y} - \overline{\boldsymbol{y}}\|^2 = \|\hat{\boldsymbol{y}} - \overline{\boldsymbol{y}}\|^2 + \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2 = SSE + SSR$, where SSE and SSR stand for the explained and residual sum of squares, respectively; R^2 is then the proportion of explained sum of squares (variation) out of the total variation, $\frac{SSE}{SST} = \frac{\|\hat{\boldsymbol{y}} - \overline{\boldsymbol{y}}\|^2}{\|\boldsymbol{y} - \overline{\boldsymbol{y}}\|^2}$. When the predictors have two groups, $\hat{\boldsymbol{y}} = \hat{\boldsymbol{x}} + \hat{\boldsymbol{z}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{Z}\hat{\boldsymbol{\gamma}}$ and $\overline{\boldsymbol{y}} = \overline{\boldsymbol{x}} + \overline{\boldsymbol{z}}$. Hence SSE can be partitioned into three parts: $SSE = \|\hat{\boldsymbol{y}} - \overline{\boldsymbol{y}}\|^2 = \|(\hat{\boldsymbol{x}} + \hat{\boldsymbol{z}}) - (\overline{\boldsymbol{x}} + \overline{\boldsymbol{z}})\|^2 = \|(\hat{\boldsymbol{x}} - \overline{\boldsymbol{x}}) + (\hat{\boldsymbol{z}} - \overline{\boldsymbol{z}})\|^2 = \|\hat{\boldsymbol{x}} - \overline{\boldsymbol{x}}\|^2 + \|\hat{\boldsymbol{z}} - \overline{\boldsymbol{z}}\|^2 + 2\langle \hat{\boldsymbol{x}} - \overline{\boldsymbol{x}}, \hat{\boldsymbol{z}} - \overline{\boldsymbol{z}} \rangle$; plugging it into the definition of R^2, we have $R^2 = \frac{SSE}{SST} = \frac{\|\hat{\boldsymbol{y}} - \overline{\boldsymbol{y}}\|^2}{\|\boldsymbol{y} - \overline{\boldsymbol{y}}\|^2} = \frac{\|\hat{\boldsymbol{x}} - \overline{\boldsymbol{x}}\|^2 + \|\hat{\boldsymbol{z}} - \overline{\boldsymbol{z}}\|^2 + 2\langle \hat{\boldsymbol{x}} - \overline{\boldsymbol{x}}, \hat{\boldsymbol{z}} - \overline{\boldsymbol{z}} \rangle}{\|\boldsymbol{y} - \overline{\boldsymbol{y}}\|^2} = \frac{SSE_x}{SST} + \frac{SSE_z}{SST} + \frac{SSE_{xz}}{SST} = R_x^2 + R_z^2 + R_{xz}^2$, which we interpret as the proportions of total variation explained by the first group of predictors, the second, and their interactions, respectively.

Supplementary References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169.

Blankenberg, D., Von Kuster, G., Coraor, N., Ananda, G., Lazarus, R., Mangan, M., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol Chapter 19, Unit 19 10 11-21.

Ercan, S., Giresi, P.G., Whittle, C.M., Zhang, X., Green, R.D., and Lieb, J.D. (2007). X chromosome repression by localization of the C. elegans dosage compensation machinery to sites of transcription initiation. Nat Genet 39, 403-408.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Y., Blankenberg, D., Albert, I., Taylor, J., Miller, W., Kent, W.J., and Nekrutenko, A. (2005). Galaxy: a platform for interactive large-scale genome analysis. Genome research 15, 1451-1455.

Goecks, J., Nekrutenko, A., Taylor, J., and Galaxy, T. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome biology 11, R86.

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C., Fisher-Aylor, K.I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A.J., Hoffman, M.M., Iyer, V.R., Jung, Y.L., Karmakar, S., Kellis, M., Kharchenko, P.V., Li, Q., Liu, T., Liu, X.S., Ma, L., Milosavljevic, A., Myers, R.M., Park, P.J., Pazin, M.J., Perry, M.D., Raha, D., Reddy, T.E., Rozowsky, J., Shoresh, N., Sidow, A., Slattery, M., Stamatoyannopoulos, J.A., Tolstorukov, M.Y., White, K.P., Xi, S., Farnham, P.J., Lieb, J.D., Wold, B.J., and Snyder, M. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res 22, 1813-1831.

Levine, M.E., Suarez, J.A., Brandhorst, S., Balasubramanian, P., Cheng, C.W., Madia, F., Fontana, L., Mirisola, M.G., Guevara-Aguirre, J., Wan, J., Passarino, G., Kennedy, B.K., Wei, M., Cohen, P., Crimmins, E.M., and Longo, V.D. (2014). Low protein intake is associated with a major reduction in IGF-1, cancer, and overall mortality in the 65 and younger but not older population. Cell metabolism 19, 407-417.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15, 550.

Nikolayeva, O., and Robinson, M.D. (2014). edgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology. Methods Mol Biol 1150, 45-79.

Oosting, J., Eilers, P., and Menezes, R. (2014). quantsmooth: Quantile smoothing and genomic visualization of array data. J. Oosting, ed., p. R package.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., Ali, S., Chin, S.F., Palmieri, C., Caldas, C., and Carroll, J.S. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481, 389-393.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics 15, 284.
Shestov, A.A., Liu, X., Ser, Z., Cluntun, A.A., Hung, Y.P., Huang, L., Kim, D., Le, A., Yellen, G., Albeck, J.G., and Locasale, J.W. (2014). Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step. eLife 3.
Stark, R., and Brown, G. (2011). DiffBind: differential binding analysis of ChIP-Seq peak data. Bioconductor.
Thalacker-Mercer, A., Stec, M., Cui, X., Cross, J., Windham, S., and Bamman, M. (2013). Cluster analysis reveals differential transcript profiles associated with resistance training-induced human skeletal muscle hypertrophy. Physiological genomics 45, 499-507.
Thalacker-Mercer, A.E., Petrella, J.K., and Bamman, M.M. (2009). Does habitual dietary intake influence myofiber hypertrophy in response to resistance training? A cluster analysis. Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme 34, 632-639.
Yu, G., Wang, L.G., and He, Q.Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics.