Supplementary Figures

Figure S1. Alternative similarity metrics. (A) The similarity of two TFs was defined as the proportion of genes affected by both RNAi experiments among all the genes affected by either RNAi experiment. (B) Pearson correlation of the NACEP distances of two RNAi experiments, restricted to the top 5% genes most strongly affected by RNAi experiments.

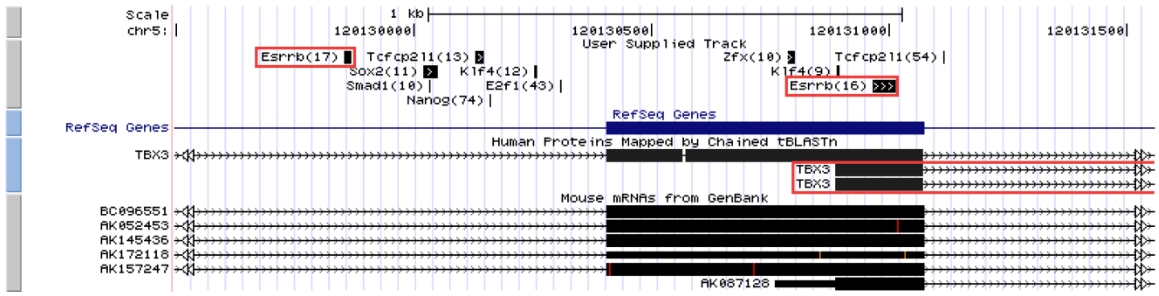
A

Top List Intersection	Nanog	Oct4	Sox2	Esrrb	Ccnblip1	Tbx3	Tcl1	RA
Nanog	1.00	0.53	0.58	0.46	0.43	0.53	0.48	0.20
Oct4	0.53	1.00	0.57	0.40	0.36	0.42	0.42	0.25
Sox2	0.58	0.57	1.00	0.55	0.38	0.48	0.43	0.28
Esrrb	0.46	0.40	0.55	1.00	0.29	0.48	0.35	0.30
Ccnblip1	0.43	0.36	0.38	0.29	1.00	0.48	0.48	0.16
Tbx3	0.53	0.42	0.48	0.48	0.48	1.00	0.53	0.29
Tcl1	0.48	0.42	0.43	0.35	0.48	0.53	1.00	0.18
RA	0.20	0.25	0.28	0.30	0.16	0.29	0.18	1.00

B

Top Genes Correlation	Nanog	Oct4	Sox2	Esrrb	Tbx3	Ccnblip1	Tcl1	RA
Nanog	1.00	0.24	0.31	0.06	-0.07	-0.09	-0.13	-0.39
Oct4	0.24	1.00	0.39	0.03	-0.15	-0.20	-0.28	-0.38
Sox2	0.31	0.39	1.00	0.32	-0.04	-0.15	-0.14	-0.34
Esrrb	0.06	0.03	0.32	1.00	0.28	-0.07	-0.07	-0.18
Tbx3	-0.07	-0.15	-0.04	0.28	1.00	0.16	0.03	-0.31
Ccnblip1	-0.09	-0.20	-0.15	-0.07	0.16	1.00	0.00	-0.46
Tcl1	-0.13	-0.28	-0.14	-0.07	0.03	0.00	1.00	-0.47
RA	-0.39	-0.38	-0.34	-0.18	-0.31	-0.46	-0.47	1.00

Figure S2. The Tbx3 gene locus on mouse chromosome 5. ChIP-seq peaks are shown in user-supplied tracks, the number of overlapping ChIP-seq reads are shown in brackets.



Supplementary text

The Gibbs Sampler algorithm for NACEP model inference.

Denote the expression data of all genes as y, the collection of all model parameters as θ, and the initial values of b, σ as $b^{(0)}$, $\sigma^{(0)}$.
For Step i in the iteration,
1. sample θ:

Sample $\theta_1^{(i)}$ from $\theta_1^{(i)}|\theta_{-(1)}^{(i-1)}, b^{(i-1)}, \sigma^{(i-1)}, y$;

Sample $\theta_2^{(i)}$ from $\theta_2^{(i)}|\theta_1^{(i)}, \theta_{-(1,2)}^{(i-1)}, b^{(i-1)}, \sigma^{(i-1)}, y$;

.

.

Sample $\theta_n^{(i)}$ from $\theta_n^{(i)}|\theta_{-(n)}^{(i)}, b^{(i-1)}, \sigma^{(i-1)}, y$;

2. Re-sample θ. This is not a typical step in Gibbs Samplers. The purpose of this step is to speed-up the convergence of the Gibbs Sample (see Bush and MacEachern (1996), [Ref.15]). We use γ to indicate re-sampled θ and suppose the number of clusters in Step i is K_i.

Sample $\gamma_1^{(i)}$ from $\gamma_1^{(i)}|b^{(i-1)}, \sigma^{(i-1)}, C^{(i-1)}$;

Sample $\gamma_2^{(i)}$ from $\gamma_2^{(i)}|b^{(i-1)}, \sigma^{(i-1)}, C^{(i-1)}$;

.

.

Sample $\gamma_{K_i}^{(i)}$ from $\gamma_{K_i}^{(i)}|b^{(i-1)}, \sigma^{(i-1)}, C^{(i-1)}$;

3. sample gene effect b:

Sample $b_1^{(i)}$ from $b_1^{(i)}|b_{-(1)}^{(i-1)}, \theta^{(i)}, \sigma^{(i-1)}, y$;

Sample $b_2^{(i)}$ from $b_2^{(i)}|b_1^{(i)}, b_{-(1,2)}^{(i-1)}, \theta^{(i)}, \sigma^{(i-1)}, y$;

.

.

Sample $b_n^{(i)}$ from $b_n^{(i)}|b_{-(n)}^{(i-1)}, \theta^{(i)}, \sigma^{(i-1)}, y$;

4. sample σ:

Sample $\sigma^{(i)}$ from $\sigma^{(i)}|\theta^{(i)}, b^{(i)}, y$;

where $\text{Parameter}_{-(k,l)}^{(i-1)}$ indicates a vector including all $\text{Parameter}^{(i-1)}$ except

$\mathrm{Parameter}_k^{(i-1)}$, $\mathrm{Parameter}_l^{(i-1)}$.

The conditional distributions are as follows. $\theta_{-k}^{(i)}$ denotes vector θ in Step i without θ_k.

1.

$$p(\theta_k^{(i)}|\theta_{-k}^{(i)}, b^{(i-1)}, \sigma^{(i-1)}, y) \propto \sum_{l\neq k} \varphi\left(y_k \middle| X\beta_l^{(i)} + b_k^{(i-1)}L, \sigma^{(i-1)2}I\right) N(0, \phi_l^{(i)2})\delta_{\theta_l^{(i)}} + \alpha H \cdot G^*,$$

where

$$H = \int \varphi\left(y_k \middle| X\beta_l^{(i)} + b_k^{(i-1)}L, \sigma^{(i)2}I\right)\varphi\left(\beta_k^{(i)} \middle| \beta_0, \sigma^{(i-1)2}(X^TX)^{-1}\right) IG\left(\phi_k^{(i)2} \middle| e, f\right) N\left(0, \phi_l^{(i)2}\right) d\beta_k^{(i)} d\phi_k^{(i)2}$$

$$= (2\pi)^{-\frac{m+1}{2}}\left(1 + \sigma^{(i-1)2}\right)^{\frac{p}{2}}\left(\sigma^{(i-1)2}\right)^{-\frac{m+p}{2}} \frac{\Gamma\left(e + \frac{1}{2}\right) f^e}{\Gamma(e)\left(f + \frac{b_k^{(i-1)2}}{2}\right)^{e+\frac{1}{2}}} \cdot$$

$$\exp\left(-\frac{1}{2\sigma^{(i-1)2}}\left(y_k - b_k^{(i)}L\right)^T\left(y_k - b_k^{(i)}L\right) - \frac{1}{2}\beta_0^T((X^TX)\beta_0 + \frac{1}{2}\left(\frac{\sigma^{(i-1)2}}{1+\sigma^{(i-1)2}}\right)\left(\frac{X^T\left(y_k - b_k^{(i-1)}L\right)}{\sigma^{(i-1)2}} + X^TX\beta_0\right)^T (X^TX)^{-1}\left(\frac{X^T\left(y_k - b_k^{(i-1)}L\right)}{\sigma^{(i-1)2}} + X^TX\beta_0\right)\right);$$

$$G^* = G^*\left(\beta_k^{(i)}\right) \cdot G^*\left(\phi_k^{(i)2}\right);$$

$$G^*\left(\beta_k^{(i)}\right) \sim N\left(\left(\frac{\sigma^{(i-1)2}}{1 + \sigma^{(i-1)2}}\right)(X^TX)^{-1}X^T\left(\frac{\left(y_k - b_k^{(i-1)}L\right)}{\sigma^{(i-1)2}} + X\beta_0\right), \frac{\sigma^{(i-1)2}}{1 + \sigma^{(i-1)2}}(X^TX)^{-1}\right);$$

$$G^*\left(\phi_k^{(i)2}\right) \sim IG\left(\phi_k^{(i)2} \middle| e + \frac{1}{2}, f + \frac{b_k^{(i-1)2}}{2}\right).$$

2.

$$p\left(\gamma_k^{(i)} \middle| b^{(i-1)}, \sigma^{(i-1)}, y\right) = p\left(\gamma_{\beta_k}^{(i)} \middle| b^{(i-1)}, \sigma^{(i-1)}, y\right) \cdot p\left(\gamma_{\phi_k^2}^{(i)} \middle| b^{(i-1)}, \sigma^{(i-1)}, y\right)$$

$$p\left(\gamma_{\beta_k}^{(i)} \middle| b^{(i-1)}, \sigma^{(i-1)}, y\right)$$

$$\sim N\left(\left(\frac{\sigma^{(i-1)2}}{n(z) + \sigma^{(i-1)2}}\right)(X^TX)^{-1}X^T\left(\frac{\sum_{z=1}^{n(z)}\left(y_k - b_k^{(i-1)}L\right)}{\sigma^{(i-1)2}} + X\beta_0\right), \frac{\sigma^{(i-1)2}}{n(z) + \sigma^{(i-1)2}}(X^TX)^{-1}\right);$$

$$p\left(\gamma_{\phi_k^2}^{(i)} \middle| b^{(i-1)}, \sigma^{(i-1)}, y\right) \sim IG\left(e + \frac{n(z)}{2}, f + \frac{1}{2}\sum_{z=1}^{n(z)} b_z^{(i-1)2}\right)$$

3.

$$p(b_k^{(i)}|b_{-k}^{(i-1)},\gamma_k^{(i)},\sigma^{(i-1)},y)\sim N(\frac{\gamma_{\phi_k^2}^{(i)2}\left(y_k-X\gamma_{\beta_k}^{(i)2}\right)^T L}{m\gamma_{\phi_k^2}^{(i)2}+\sigma^{(i-1)2}},\frac{\sigma^{(i-1)2}(X^TX)^{-1}}{m\gamma_{\phi_k^2}^{(i)2}+\sigma^{(i-1)2}})$$

4.

$$p\left(\sigma^{(i)2}\middle|\gamma^{(i)},b^{(i)},y\right)\sim IG\left(g+\frac{nm}{2},h+\frac{1}{2}\sum_{z=1}^{K}\sum_{l=1}^{n(z)}\left(Y_{zl}-X\gamma_{\beta_z}^{(i)}-b_{zl}^{(i)}L\right)^T\left(Y_{zl}-X\gamma_{\beta_z}^{(i)}-b_{zl}^{(i)}L\right)\right).$$