A Bayesian model for hybrid vigor

Will Landau

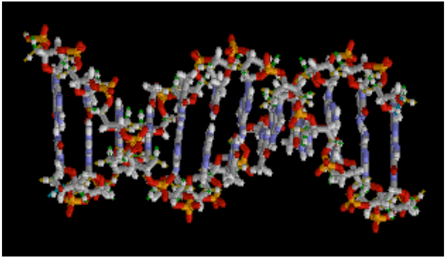
Iowa State University

September 29, 2013

Outline

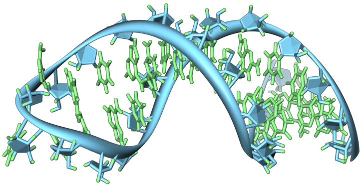
Outline

DNA



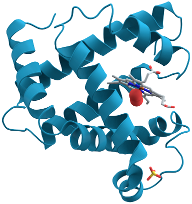
...GTGCATCTGACTCCTGAGGAGAAG...
...CACGTAGACTGAGGACTCCTCTTC...

RNA



... GUGCAUCUGACUCCUGAGGAGAAG ...

Proteins



··· V — H — L — T — P — E — E — K ···

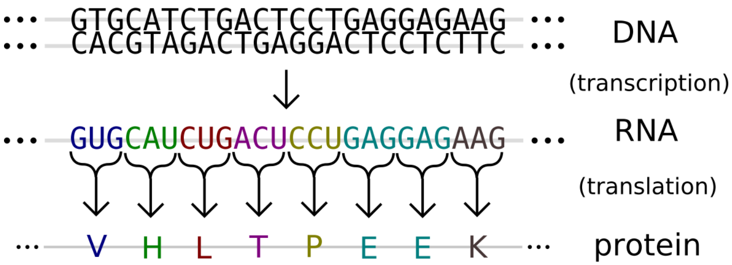
Central dogma: how organisms make proteins

... GTGCATCTGACTCCTGAGGAGAAG ...
... CACGTAGACTGAGGACTCCTCTTC ...

DNA

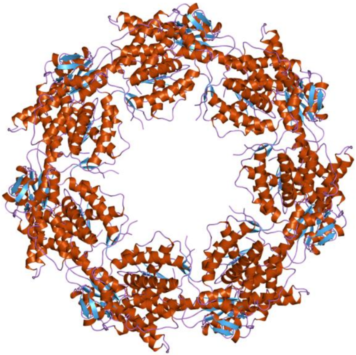
Central dogma: how organisms make proteins

Central dogma of genetics



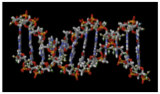
HSP60

- HSP = heat shock protein.
- Prevent heat damage to other proteins.

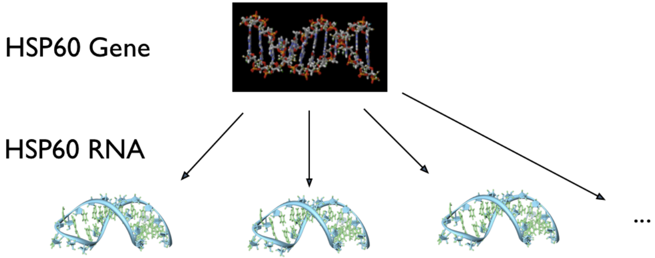


Temperature spike triggers HSP60 production.

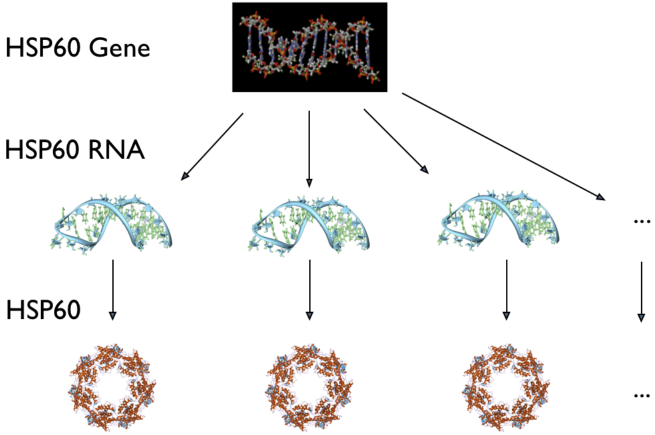
HSP60 Gene



Temperature spike causes HSP60 expression.

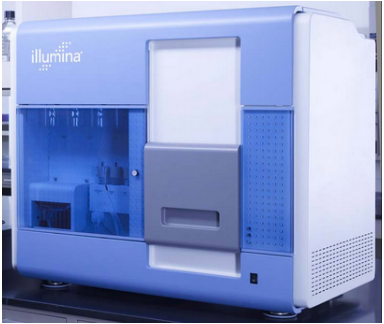


Temperature spike causes HSP60 expression.

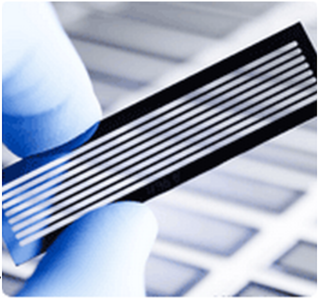


RNA-seq

- RNA sequencing: measure gene expression using relative abundance of RNA.
- Illumina Genome Analyzer:



/

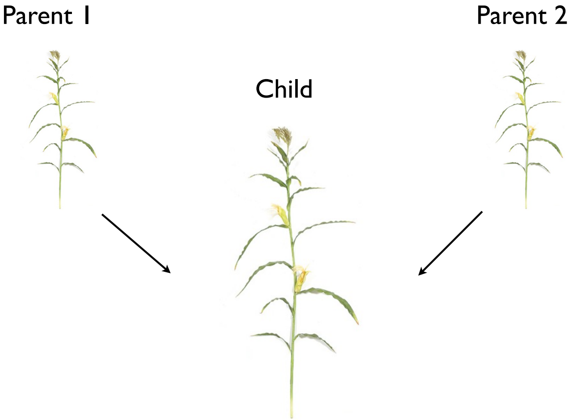


RNA-seq data: counts of amplified RNA fragments

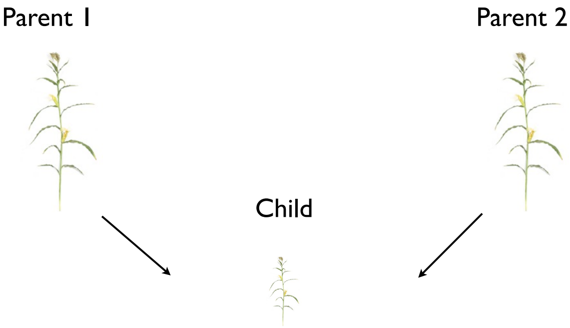
	Treatment 1			Treatment 2		Treatment 3	
Gene 1	100	225	0	70	279	300	106
Gene 2	0	1	1	50	501	2	7
Gene 3	3	4	2	700	900	0	0
Gene 4	893	400	760	5	5	1000	513
...	...	...	...	...	...	...	...
Gene 34897	10	13	6	819	761	902	912

- Goal: use RNA-seq to study hybrid vigor (heterosis).

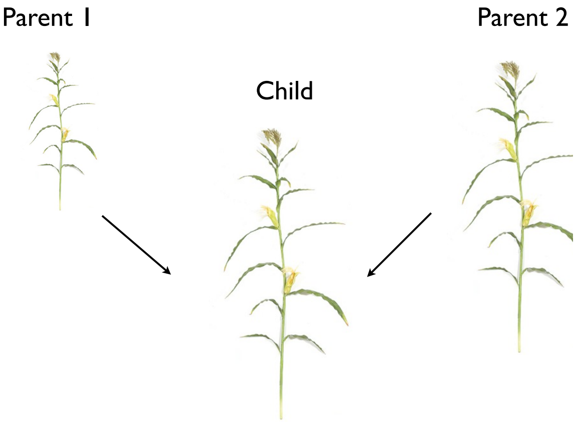
High-parent heterosis: child's trait surpasses both parents



Low-parent heterosis: child's trait is weaker than in each parent



Mid-parent heterosis: child's trait is different than average of parents



High-parent heterosis in gene expression

	Parent 1			Child		Parent 2	
Gene 1	100	225	0	70	279	300	106
Gene 2	0	1	1	50	501	2	7
Gene 3	3	4	2	700	900	0	0
Gene 4	893	400	760	5	5	1000	513
...	...	...	...	...	...	...	...
Gene 34897	10	13	6	819	761	902	912

Low-parent heterosis in gene expression

	Parent 1			Child		Parent 2	
Gene 1	100	225	0	70	279	300	106
Gene 2	0	1	1	50	501	2	7
Gene 3	3	4	2	700	900	0	0
Gene 4	893	400	760	5	5	1000	513
...	...	...	...	...	...	...	...
Gene 34897	10	13	6	819	761	902	912

Mid-parent heterosis in gene expression

	Parent 1			Child		Parent 2	
Gene 1	100	225	0	70	279	300	106
Gene 2	0	1	1	50	501	2	7
Gene 3	3	4	2	700	900	0	0
Gene 4	893	400	760	5	5	1000	513
...	...	...	...	...	...	...	...
Gene 34897	10	13	6	819	761	902	912

Outline

The model

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

The model

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$c_n \stackrel{\text{ind}}{\sim} \text{N}(c_n \mid 0, \sigma_c^2)$$

The model

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$c_n \overset{\text{ind}}{\sim} \text{N}(c_n \mid 0, \sigma_c^2)$$

$$\varepsilon_{g,n} \overset{\text{ind}}{\sim} \text{N}(\varepsilon_{g,n} \mid 0, \eta_g^2)$$

The model

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$c_n \overset{\text{ind}}{\sim} \text{N}(c_n \mid 0, \sigma_c^2)$$

$$\sigma_c \sim \text{U}(\sigma_c \mid 0, \sigma_{c0})$$

$$\varepsilon_{g,n} \overset{\text{ind}}{\sim} \text{N}(\varepsilon_{g,n} \mid 0, \eta_g^2)$$

The model

$$
\begin{aligned}
y_{g,n} &\stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \\
c_n &\stackrel{\text{ind}}{\sim} \text{N}(c_n \mid 0, \sigma_c^2) \\
\sigma_c &\sim \text{U}(\sigma_c \mid 0, \sigma_{c0}) \\
\varepsilon_{g,n} &\stackrel{\text{ind}}{\sim} \text{N}(\varepsilon_{g,n} \mid 0, \eta_g^2) \\
\eta_g^2 &\stackrel{\text{ind}}{\sim} \text{Inv-Gamma}\left(\eta_g^2 \,\middle|\, \text{shape} = \frac{d}{2}, \ \text{rate} = \frac{d \cdot \tau^2}{2}\right)
\end{aligned}
$$

The model

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$c_n \stackrel{\text{ind}}{\sim} \text{N}(c_n \mid 0, \sigma_c^2)$$

$$\sigma_c \sim \text{U}(\sigma_c \mid 0, \sigma_{c0})$$

$$\varepsilon_{g,n} \stackrel{\text{ind}}{\sim} \text{N}(\varepsilon_{g,n} \mid 0, \eta_g^2)$$

$$\eta_g^2 \stackrel{\text{ind}}{\sim} \text{Inv-Gamma}\left(\eta_g^2 \,\middle|\, \text{shape} = \frac{d}{2},\ \text{rate} = \frac{d \cdot \tau^2}{2}\right)$$

$$d \sim \text{U}(d \mid 0, d_0)$$

The model

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$c_n \overset{\text{ind}}{\sim} \text{N}(c_n \mid 0, \sigma_c^2)$$

$$\sigma_c \sim \text{U}(\sigma_c \mid 0, \sigma_{c0})$$

$$\varepsilon_{g,n} \overset{\text{ind}}{\sim} \text{N}(\varepsilon_{g,n} \mid 0, \eta_g^2)$$

$$\eta_g^2 \overset{\text{ind}}{\sim} \text{Inv-Gamma}\left(\eta_g^2 \,\middle|\, \text{shape} = \frac{d}{2}, \ \text{rate} = \frac{d \cdot \tau^2}{2}\right)$$

$$d \sim \text{U}(d \mid 0, d_0)$$

$$\tau^2 \sim \text{Gamma}(\tau^2 \mid \text{shape} = a_\tau, \text{rate} = b_\tau)$$

The model

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma^2_\phi)$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\theta_\alpha \sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\theta_\alpha \sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

$$\theta_\delta \sim \text{N}(\theta_\delta \mid 0, \gamma_\delta^2)$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\sigma_\phi \sim \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0})$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\theta_\alpha \sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

$$\theta_\delta \sim \text{N}(\theta_\delta \mid 0, \gamma_\delta^2)$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\sigma_\phi \sim \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0})$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\theta_\alpha \sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)$$

$$\sigma_\alpha \sim \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0})$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

$$\theta_\delta \sim \text{N}(\theta_\delta \mid 0, \gamma_\delta^2)$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\sigma_\phi \sim \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0})$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\theta_\alpha \sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)$$

$$\sigma_\alpha \sim \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0})$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

$$\theta_\delta \sim \text{N}(\theta_\delta \mid 0, \gamma_\delta^2)$$

$$\sigma_\delta \sim \text{U}(\sigma_\delta \mid 0, \sigma_{\delta 0})$$

The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$\phi_g \overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2)$$

$$\theta_\phi \sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2)$$

$$\sigma_\phi \sim \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0})$$

$$\alpha_g \overset{\text{ind}}{\sim} \pi_\alpha^{1 - I(\alpha_g)}[(1 - \pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)}$$

$$\theta_\alpha \sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2)$$

$$\sigma_\alpha \sim \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0})$$

$$\pi_\alpha \sim \text{Beta}(\pi_\alpha \mid a_\alpha, b_\alpha)$$

$$\delta_g \overset{\text{ind}}{\sim} \pi_\delta^{1 - I(\delta_g)}[(1 - \pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)}$$

$$\theta_\delta \sim \text{N}(\theta_\delta \mid 0, \gamma_\delta^2)$$

$$\sigma_\delta \sim \text{U}(\sigma_\delta \mid 0, \sigma_{\delta 0})$$

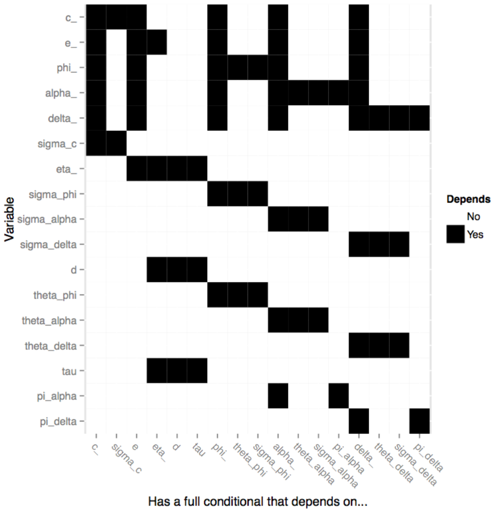
The model

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$\begin{aligned}
y_{g,n} &\overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g))) \\
\phi_g &\overset{\text{ind}}{\sim} \text{N}(\phi_g \mid \theta_\phi, \sigma_\phi^2) \\
\theta_\phi &\sim \text{N}(\theta_\phi \mid 0, \gamma_\phi^2) \\
\sigma_\phi &\sim \text{U}(\sigma_\phi \mid 0, \sigma_{\phi 0}) \\
\alpha_g &\overset{\text{ind}}{\sim} \pi_\alpha^{1-I(\alpha_g)}[(1-\pi_\alpha)\text{N}(\alpha_g \mid \theta_\alpha, \sigma_\alpha^2)]^{I(\alpha_g)} \\
\theta_\alpha &\sim \text{N}(\theta_\alpha \mid 0, \gamma_\alpha^2) \\
\sigma_\alpha &\sim \text{U}(\sigma_\alpha \mid 0, \sigma_{\alpha 0}) \\
\pi_\alpha &\sim \text{Beta}(\pi_\alpha \mid a_\alpha, b_\alpha) \\
\delta_g &\overset{\text{ind}}{\sim} \pi_\delta^{1-I(\delta_g)}[(1-\pi_\delta)\text{N}(\delta_g \mid \theta_\delta, \sigma_\delta^2)]^{I(\delta_g)} \\
\theta_\delta &\sim \text{N}(\theta_\delta \mid 0, \gamma_\delta^2) \\
\sigma_\delta &\sim \text{U}(\sigma_\delta \mid 0, \sigma_{\delta 0}) \\
\pi_\delta &\sim \text{Beta}(\pi_\delta \mid a_\delta, b_\delta)
\end{aligned}$$

Outline

Partition parameters by conditional independence.



Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ \ldots, \ c_N$

Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ \ldots, \ c_N$
2. $\tau, \ \pi_\alpha, \ \pi_\delta$

Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ \ldots, \ c_N$
2. $\tau, \ \pi_\alpha, \ \pi_\delta$
3. $d, \ \theta_\phi, \ \theta_\alpha, \ \theta_\delta$

Use these partitions as Gibbs steps.

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ \ldots, \ c_N$
2. $\tau, \ \pi_\alpha, \ \pi_\delta$
3. $d, \ \theta_\phi, \ \theta_\alpha, \ \theta_\delta$
4. $\sigma_c, \ \sigma_\phi, \ \sigma_\alpha, \ \sigma_\delta, \ \eta_1^2, \ \ldots, \ \eta_G^2$

Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ldots, c_N$
2. $\tau, \pi_\alpha, \pi_\delta$
3. $d, \theta_\phi, \theta_\alpha, \theta_\delta$
4. $\sigma_c, \sigma_\phi, \sigma_\alpha, \sigma_\delta, \eta_1^2, \ldots, \eta_G^2$
5. $\varepsilon_{1,1}, \varepsilon_{1,2}, \ldots, \varepsilon_{1,N}, \varepsilon_{2,N}, \ldots, \varepsilon_{G,N}$

Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ldots, c_N$
2. $\tau, \pi_\alpha, \pi_\delta$
3. $d, \theta_\phi, \theta_\alpha, \theta_\delta$
4. $\sigma_c, \sigma_\phi, \sigma_\alpha, \sigma_\delta, \eta_1^2, \ldots, \eta_G^2$
5. $\varepsilon_{1,1}, \varepsilon_{1,2}, \ldots, \varepsilon_{1,N}, \varepsilon_{2,N}, \ldots, \varepsilon_{G,N}$
6. $\phi_1, \ldots, \phi_G$

Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ \ldots, \ c_N$
2. $\tau, \ \pi_\alpha, \ \pi_\delta$
3. $d, \ \theta_\phi, \ \theta_\alpha, \ \theta_\delta$
4. $\sigma_c, \ \sigma_\phi, \ \sigma_\alpha, \ \sigma_\delta, \ \eta_1^2, \ \ldots, \ \eta_G^2$
5. $\varepsilon_{1,1}, \ \varepsilon_{1,2}, \ \ldots, \ \varepsilon_{1,N}, \ \varepsilon_{2,N}, \ \ldots, \ \varepsilon_{G,N}$
6. $\phi_1, \ \ldots, \ \phi_G$
7. $\alpha_1, \ \ldots, \ \alpha_G$

Use these partitions as Gibbs steps.

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- From the appropriate full conditional distributions, sample the following:

1. $c_1, \ldots, c_N$
2. $\tau, \pi_\alpha, \pi_\delta$
3. $d, \theta_\phi, \theta_\alpha, \theta_\delta$
4. $\sigma_c, \sigma_\phi, \sigma_\alpha, \sigma_\delta, \eta_1^2, \ldots, \eta_G^2$
5. $\varepsilon_{1,1}, \varepsilon_{1,2}, \ldots, \varepsilon_{1,N}, \varepsilon_{2,N}, \ldots, \varepsilon_{G,N}$
6. $\phi_1, \ldots, \phi_G$
7. $\alpha_1, \ldots, \alpha_G$
8. $\delta_1, \ldots, \delta_G$

- and then repeat.

Estimated heterosis probabilities

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

$$I(\delta_g^{(i)} > |\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} < -|\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} \neq 0)$$

Estimated heterosis probabilities

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

Consider one chain with M iterations.

$$I(\delta_g^{(i)} > |\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} < -|\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} \neq 0)$$

Estimated heterosis probabilities

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

Consider one chain with M iterations.

$$P(\text{high-parent heterosis in gene } g) \approx \frac{1}{M} \sum_{i=1}^{M} I(\delta_g^{(i)} > |\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} < -|\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} \neq 0)$$

Estimated heterosis probabilities

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

Consider one chain with M iterations.

$$P(\text{high-parent heterosis in gene } g) \approx \frac{1}{M} \sum_{i=1}^{M} I(\delta_g^{(i)} > |\alpha_g^{(i)}|)$$

$$P(\text{low-parent heterosis in gene } g\ ) \approx \frac{1}{M} \sum_{i=1}^{M} I(\delta_g^{(i)} < -|\alpha_g^{(i)}|)$$

$$I(\delta_g^{(i)} \neq 0)$$

Estimated heterosis probabilities

$$\mu(n, \phi_g, \alpha_g, \delta_g) = \begin{cases} \phi_g - \alpha_g & \text{sample } n \text{ from parent 1} \\ \phi_g + \delta_g & \text{sample } n \text{ from child} \\ \phi_g + \alpha_g & \text{sample } n \text{ from parent 3} \end{cases}$$

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

Consider one chain with M iterations.

$$P(\text{high-parent heterosis in gene } g) \approx \frac{1}{M} \sum_{i=1}^{M} I(\delta_g^{(i)} > |\alpha_g^{(i)}|)$$

$$P(\text{low-parent heterosis in gene } g\ ) \approx \frac{1}{M} \sum_{i=1}^{M} I(\delta_g^{(i)} < -|\alpha_g^{(i)}|)$$

$$P(\text{mid-parent heterosis in gene } g\ ) \approx \frac{1}{M} \sum_{i=1}^{M} I(\delta_g^{(i)} \neq 0)$$

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
    - ϕ_g's
    - α_g's
    - δ_g's

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's
- Use parallel reductions to calculate sufficient statistics for:

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's
- Use parallel reductions to calculate sufficient statistics for:
  - c_n's

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's
- Use parallel reductions to calculate sufficient statistics for:
  - c_n's
  - τ, d

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's
- Use parallel reductions to calculate sufficient statistics for:
  - c_n's
  - τ, d
  - θ_ϕ, θ_α, θ_δ

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \overset{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's
- Use parallel reductions to calculate sufficient statistics for:
  - c_n's
  - τ, d
  - θ_ϕ, θ_α, θ_δ
  - σ_ϕ, σ_α, σ_δ, σ_c

Tons of opportunity for GPU parallelism across genes!

$$y_{g,n} \stackrel{\text{ind}}{\sim} \text{Poisson}(y_{g,n} \mid \exp(c_n + \varepsilon_{g,n} + \mu(n, \phi_g, \alpha_g, \delta_g)))$$

- Sample in parallel:
  - ϕ_g's
  - α_g's
  - δ_g's
  - $\varepsilon_{g,n}$'s
  - η_g's
- Use parallel reductions to calculate sufficient statistics for:
  - c_n's
  - τ, d
  - θ_ϕ, θ_α, θ_δ
  - σ_ϕ, σ_α, σ_δ, σ_c
  - π_α, π_δ

Outline

The software

- Ordinary C and GPU-accelerated versions, along with an R package wrapper, are available for download at `https://github.com/wlandau/heterosis`.

The software

- Ordinary C and GPU-accelerated versions, along with an R package wrapper, are available for download at https://github.com/wlandau/heterosis.
- Time for a demo...

Sources

1. Some source