Complex Phenotypes

Robert Brown

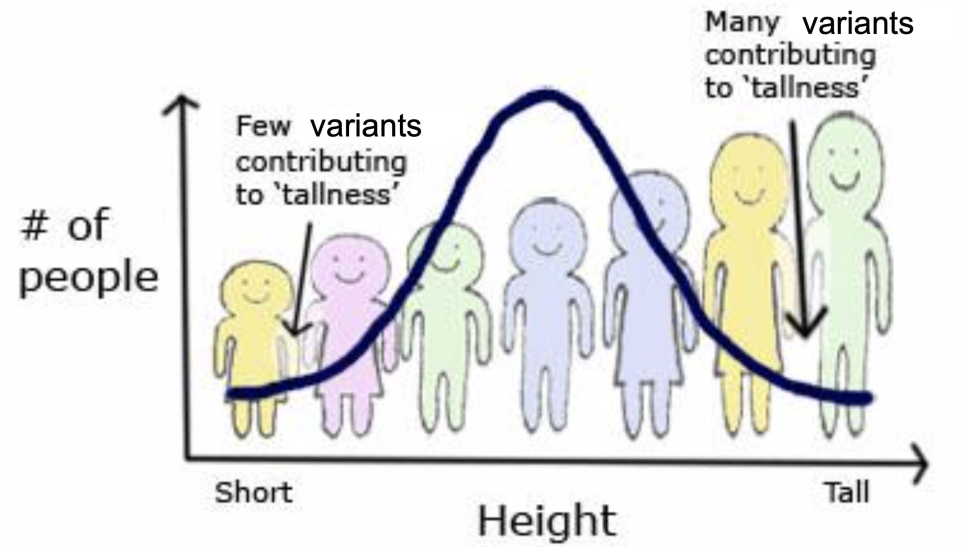
What is a complex phenotype

- Height
- BMI
- hair color
- diabetes
- autism
- annual salary
- favorite color
- blood sugar level
- RELN protein level

$$Y = Genetics + Environment$$

$$Y = G + E$$

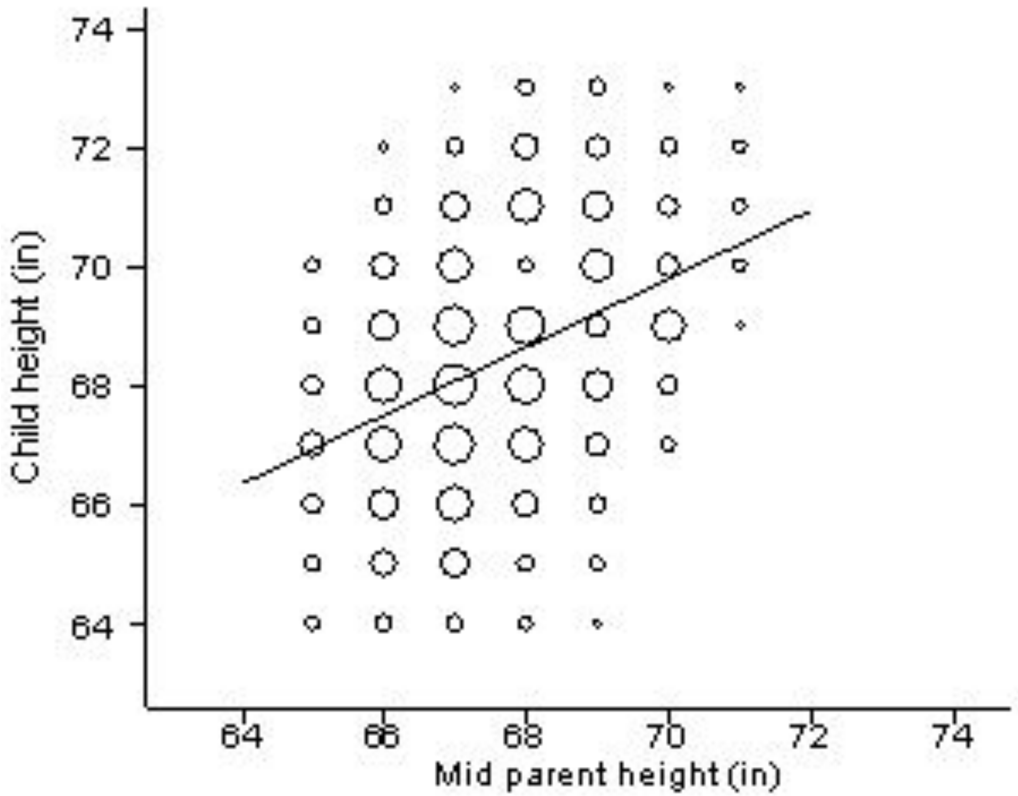
Complex phenotypes have many genetic variants contributing to their genetic component.



Complex phenotypes are "heritable"

$$Y = G + E$$

$$Y = G^p + G^m + E$$



Galton, Francis. "Regression towards mediocrity in hereditary stature." *The Journal of the Anthropological Institute of Great Britain and Ireland* 15 (1886): 246-263.

$$Y = Genetics + Environment$$

...A...T...T...C...

...G...C...T...C...

– ↓ – ↓

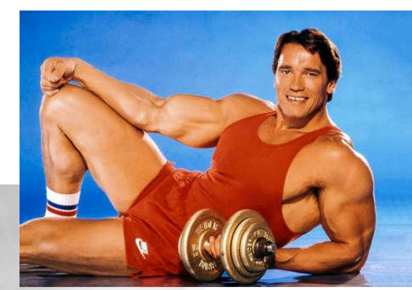
↑ – – ↓

...G...T...A...G...

...G...C...A...G...

↑ ↓ ↑ –

↑ – ↑ –



$$Y = Genetics + Environment$$

...A...T...T...C...

...G...C...T...C...

– ↓ – ↓

↑ – – ↓

...G...T...A...G...

...G...C...A...G...

↑ ↓ ↑ –

↑ – ↑ –

$$Y = G + E$$

<u>Environmental Effects</u>

- Exercise
- Diet/nutrition
- Race/ethnicity
- sex/gender
- Socio-economic status
- Parental Education
- Kindergarten Attendance
- Access to medical care
- Access to opportunity
- Distance to major cities

$$Y = Genetics + Environment$$

...A...T...T...C...

...G...C...T...C...

– ↓ – ↓

↑ – – ↓

...G...T...A...G...

...G...C...A...G...

↑ ↓ ↑ –

↑ – ↑ –

$$y = \mu + \beta_1 g_1 + \beta_2 g_2 + \beta_3 g_3 + \ldots + \beta_M g_M + e$$

$$y = \mu + 0.2 g_1 + (-0.1) g_2 + 0.3 g_3 + \ldots + (-0.2) g_M + e$$

DeVito

$$y = \mu + 0.2(1) + (-0.1)(1) + 0.3(0) + \ldots + (-0.2)(2) + e = \mu + (-0.3) + e$$

Schwarzenegger

$$y = \mu + 0.2(2) + (-0.1)(1) + 0.3(2) + \ldots + (-0.2)(0) + e = \mu + 0.9 + e$$

$$Y = Genetics + Environment$$

...A...T...T...C...

...G...C...T...C...

– ↓ – ↓

↑ – – ↓

$$y_D = \sum_i \beta_i g_{Di} + e_D$$

...G...T...A...G...

...G...C...A...G...

↑ ↓ ↑ –

↑ – ↑ –

$$y_S = \sum_i \beta_i g_{Si} + e_S$$

$$y_j = \sum_i \beta_i g_{ji} + e_j$$

DeVito

$$y = \mu + 0.2(1) + (-0.1)(1) + 0.3(0) + \ldots + (-0.2)(2) + e = \mu + (-0.3) + e$$

Schwarzenegger

$$y = \mu + 0.2(2) + (-0.1)(1) + 0.3(2) + \ldots + (-0.2)(0) + e = \mu + 0.9 + e$$

$$Y = Genetics + Environment$$

...A...T...T...C...

...G...C...T...C...

– ↓ – ↓

↑ – – ↓

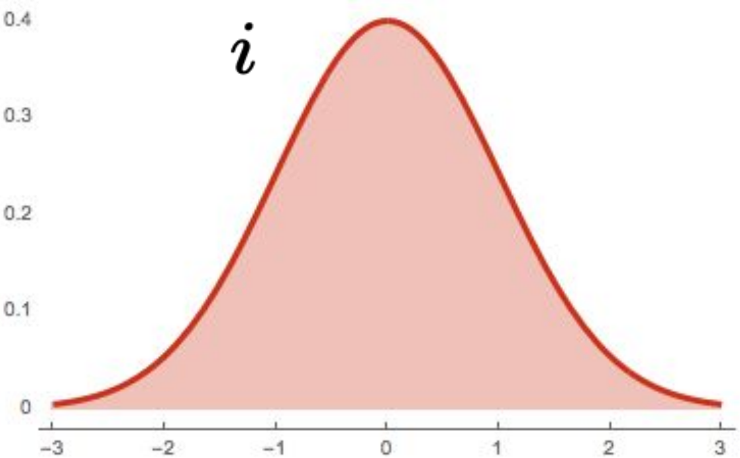
...G...T...A...G...

...G...C...A...G...

↑ ↓ ↑ –

↑ – ↑ –

$$y_j = \sum_i \beta_i g_{ji} + e_j$$



Environmental Effects

- Exercise
- Diet/nutrition
- Race/ethnicity
- sex/gender
- Socio-economic status
- Parental Education
- Kindergarten Attendance
- Access to medical care
- Access to opportunity
- Distance to major cities

$$e_j \sim \mathcal{N}(0,\, \sigma_e^2)$$

https://brilliant.org/wiki/normal-distribution/

Linear Algebra Version

$$y_j = \sum_i \beta_i g_{ji} + e_j$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1M} \\ g_{21} & g_{22} & \cdots & g_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ g_{N1} & g_{1N} & \cdots & g_{NM} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_M \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

$$Y = G\beta + E$$

Linear Algebra Version

$$y_j = \sum_i \beta_i g_{ji} + e_j$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1M} \\ g_{21} & g_{22} & \cdots & g_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ g_{N1} & g_{1N} & \cdots & g_{NM} \end{bmatrix} \begin{bmatrix} 0.2 \\ 0 \\ \vdots \\ -0.05 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

$$Y = G\beta + E$$

Types of Models

$$Y = G + E$$

Additive Genetics + Environment

$$Y = G + E + GE$$

Additive Genetics + Environment +
Gene by Environment Interaction

$$Y = G + E + GE + GG$$

Additive Genetics + Environment +
Gene by Environment Interaction +
Gene by Gene Interaction

$$y = \mu + \sum_i \beta_i g_i + e$$

$$y = \mu + \sum_i \beta_i^g g_i + \sum_i \beta_i^{gd} e^d g_i + \sum_i \beta_i^{gs} e^s g_i + \sum_i \sum_j \beta_{ij}^{gg} g_i g_j + e$$