Multivariate Genetic Analysis (Introduction)

Frühling Rijsdijk

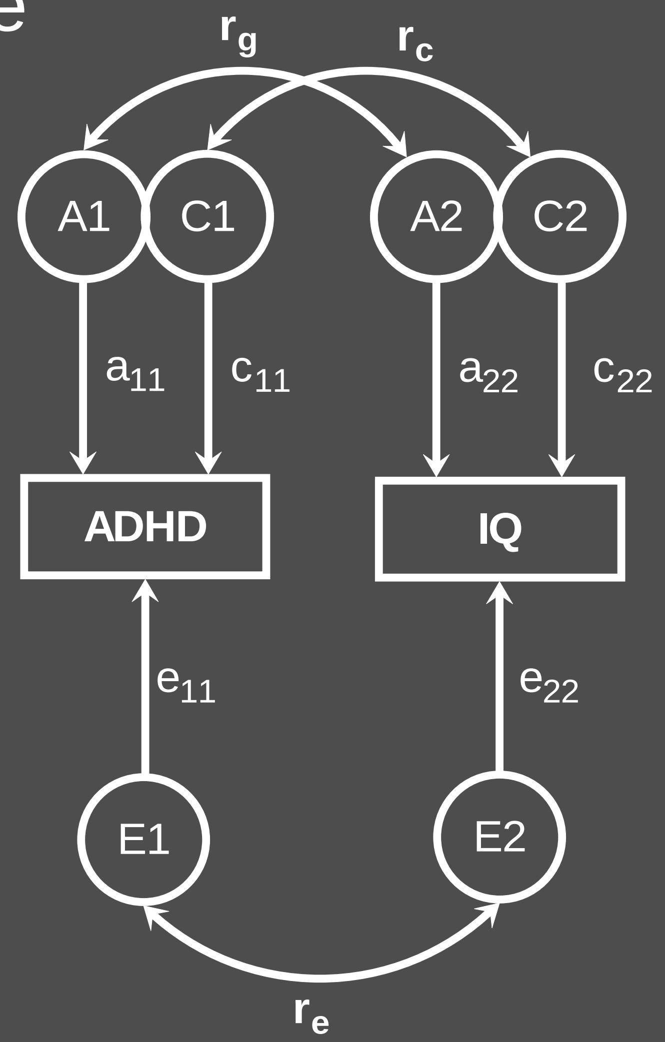
Wednesday March 8, 2006

Multivariate Twin Analyses

- Goal: to understand what factors make sets of variables correlate or co-vary
- Two or more traits can be correlated because they share common genetic or common environmental influences (C or E)
- With twin data on multiple traits it's possible to partition the covariation into it's genetic and environmental components

Example 1

Interested in reason for covariance / correlation between phenotypes, e.g. IQ and ADHD



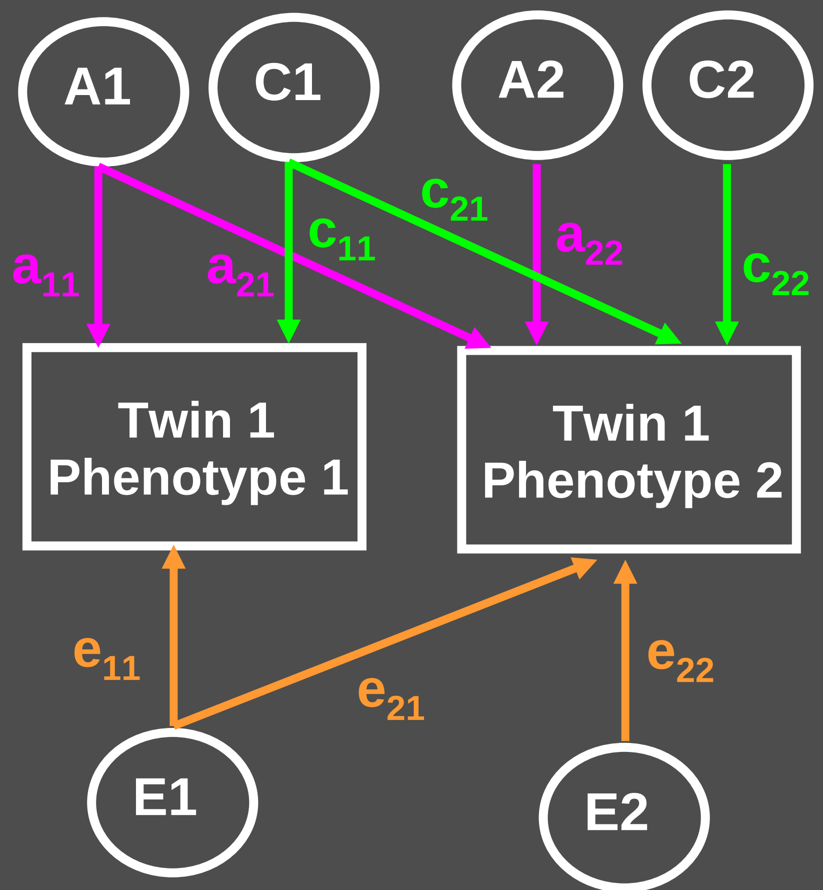
How can we explain the association?

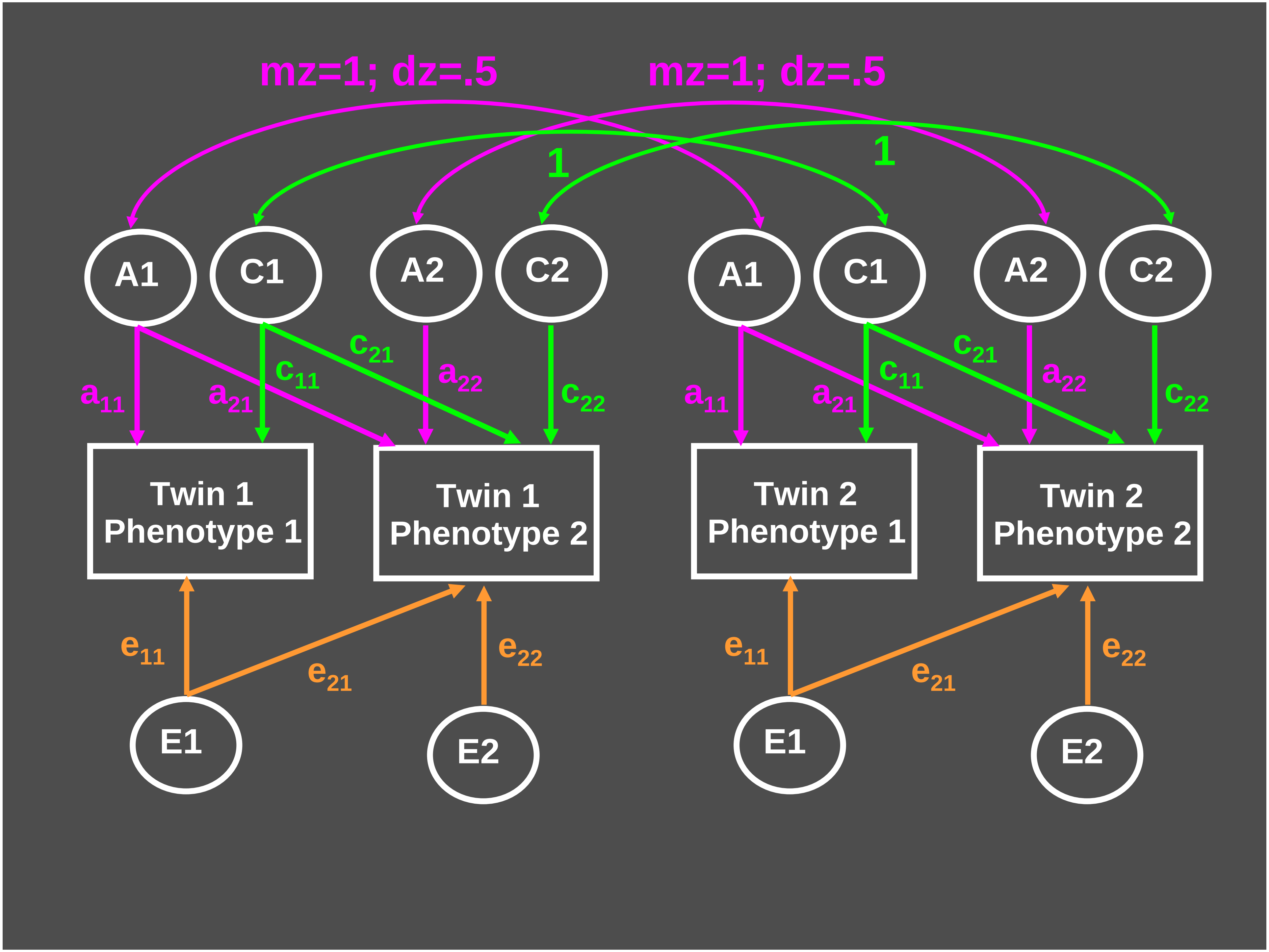
- Additive Genetic effects (r_g)
- Shared environment (r_c)
- Non-shared environment (r_e)

Kuntsi et al. *Neuropsychiatric Genetics,*

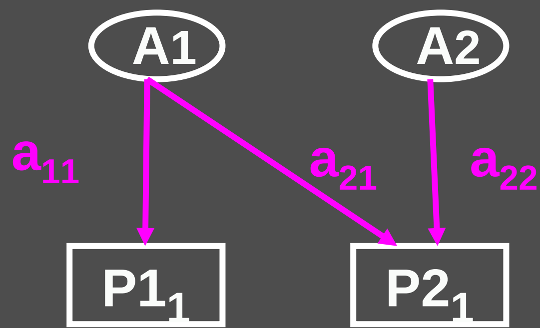
Observed Cov matrices: 4×4

		Twin1		Twin2	
		p1	p2	p1	p2
		Within-Twin Covariances			
Twin1	p1	Var P1			
	p2	Cov P1- P2	Var P2		
		Cross-Twin Covariances		Within-Twin Covariances	
Twin2	p1	Within P1	Cross Traits	Var P1	
	p2	Cross Traits	Within P2	Cov P1- P2	Var P2


A1
C1
A2
C2
a_{11}
a_{21}
c_{11}
c_{21}
a_{22}
c_{22}
Twin 1 Phenotype 1
Twin 1 Phenotype 2
e_{11}
e_{21}
e_{22}
E1
E2


mz=1; dz=.5
mz=1; dz=.5
1
1
A1
C1
A2
C2
A1
C1
A2
C2
a11
a21
c11
c21
a22
c22
a11
a21
c11
c21
a22
c22
Twin 1 Phenotype 1
Twin 1 Phenotype 2
Twin 2 Phenotype 1
Twin 2 Phenotype 2
e11
e21
e22
e11
e21
e22
E1
E2
E1
E2

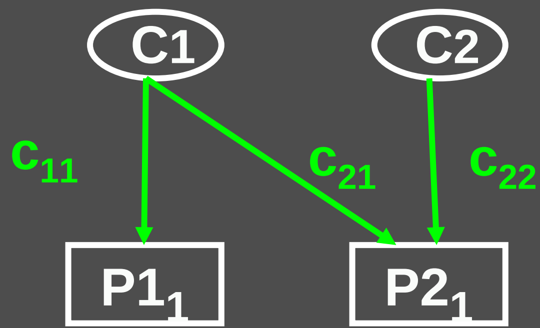
Cholesky Decomposition: Path Tracing



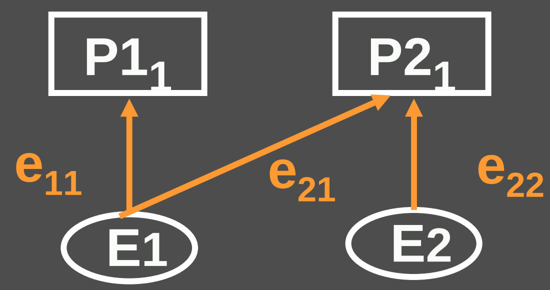
Twin1

Within-Twin Covariances

Twin1		p1	p2
	P1	a_{11}^2	
	P2	$a_{11}a_{21}$	$a_{22}^2 + a_{21}^2$

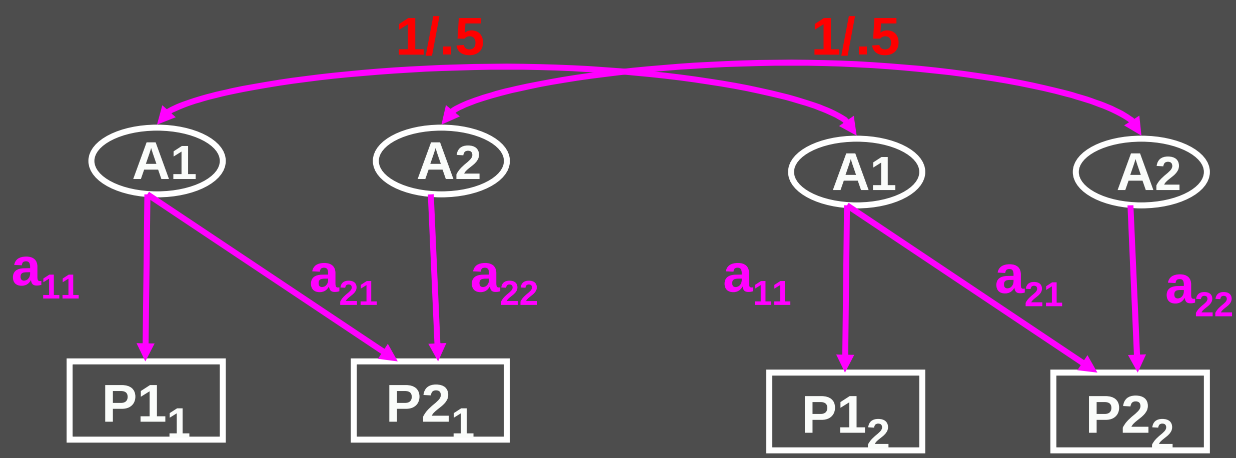


		Twin1	
		p1	p2
		Within-Twin Covariances	
Twin1	P1	$a_{11}^2 + c_{11}^2$	
	P2	$a_{11}a_{21} + c_{11}c_{21}$	$a_{22}^2 + a_{21}^2 + c_{22}^2 + c_{21}^2$



Twin1

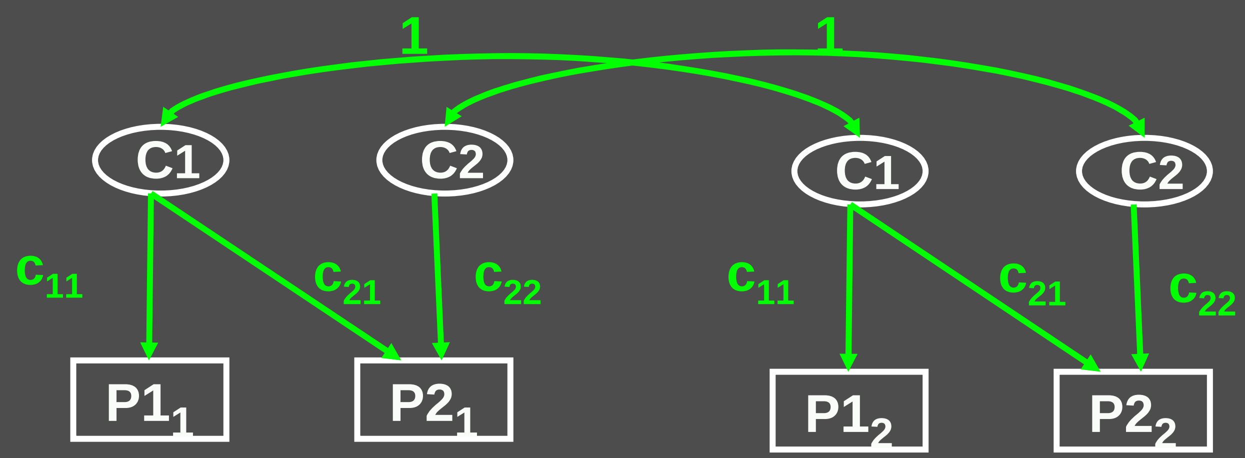
		p1	p2
		Within-Twin Covariances	
Twin1	P1	$a_{11}^2 + c_{11}^2 + e_{11}^2$	
	P2	$a_{11}a_{21} + c_{11}c_{21} + e_{11}e_{21}$	$a_{22}^2 + a_{21}^2 + c_{22}^2 + c_{21}^2 + e_{22}^2 + e_{21}^2$



Twin1

Cross-Twin Covariances

Twin2	p1	p2
P1	$1/.5a_{11}^2$	
P2	$1/.5a_{11}a_{21}$	$1/.5a_{22}^2 + 1/.5a_{21}^2$



Twin1

Cross-Twin Covariances

Twin2	p1	p2
P1	$1/.5a_{11}^2 + c_{11}^2$	
P2	$1/.5a_{11}a_{21} + c_{11}c_{21}$	$1/.5a_{22}^2 + 1/.5a_{21}^2 + c_{22}^2 + c_{21}^2$

Predicted Model

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin1	p1	Var P1			
	p2	Cov P1- P2	Var P2		
		Cross-Twin Covariances		Within-Twin Covariances	
Twin2	p1	Within P1		Var P1	
	p2	Cross Traits	Within P2	Cov P1- P2	Var P2

Var of P1 and P2 same across twins and zygosity groups

Predicted Model

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin1	p1	Var P1			
	p2	Cov P1- P2	Var P2		
		Cross-Twin Covariances		**Within-Twin Covariances**	
Twin2	p1	Within Trait 1		Var P1	
	p2	Cross Traits	Within Trait 2	Cov P1- P2	Var P2

Cov P1 - P2 same across twins and zygosity groups

Predicted Model

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin1	p1	Var P1			
	p2	Cov P1- P2	Var P2		
		Cross-Twin Covariances		**Within-Twin Covariances**	
Twin2	p1	Within P1		Var P1	
	p2	Cross Traits	Within P2	Cov P1- P2	Var P2

Cross Twin Cov within each trait, different for MZ and DZ

Predicted Model

		Twin1		Twin2	
		p1	p2	p1	p2
		Within-Twin Covariances			
Twin1	p1	Var P1			
	p2	Cov P1- P2	Var P2		
		Cross-Twin Covariances		**Within-Twin Covariances**	
Twin2	p1	Within P1		Var P1	
	p2	Cross Traits	Within P2	Cov P1- P2	Var P2

Cross Twin - Cross trait, different for MZ and DZ

MZ

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin 1	p1	1			
	p2	.30	1		
		Cross-Twin Covariances		Within-Twin Covariances	
Twin 2	p1	. 79	.49	1	
	p2	.50	.59	. 29	1

DZ

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin 1	p1	1			
	p2	.30	1		
		Cross-Twin Covariances		Within-Twin Covariances	
Twin 2	p1	.39	.25	1	
	p2	.24	.43	. 31	1

MZ

Within-Twin Covariances (Twin1 × Twin1); Cross-Twin Covariances (Twin2 × Twin1); Within-Twin Covariances (Twin2 × Twin2)

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
Twin 1	p1	1			
	p2	.30	1		
Twin 2	p1	. 79	.24	1	
	p2	.25	.59	. 29	1

DZ

Within-Twin Covariances (Twin1 × Twin1); Cross-Twin Covariances (Twin2 × Twin1); Within-Twin Covariances (Twin2 × Twin2)

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
Twin 1	p1	1			
	p2	.30	1		
Twin 2	p1	.39	.23	1	
	p2	.24	.43	. 31	1

MZ

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin 1	p1	1			
Twin 1	p2	.30	1		
		Cross-Twin Covariances		Within-Twin Covariances	
Twin 2	p1	. 79	.01	1	
Twin 2	p2	.01	.59	. 29	1

DZ

		Twin1 p1	Twin1 p2	Twin2 p1	Twin2 p2
		Within-Twin Covariances			
Twin 1	p1	1			
Twin 1	p2	.30	1		
		Cross-Twin Covariances		Within-Twin Covariances	
Twin 2	p1	.39	.01	1	
Twin 2	p2	.01	.43	. 31	1

Summary

- Within-individual cross-trait covariance implies common etiological influences
- Cross-twin cross-trait covariance implies that these common etiological influences are familial
- Whether these common familial influences are genetic or environmental, is reflected in the MZ/DZ ratio of the cross-twin cross-traits covariances

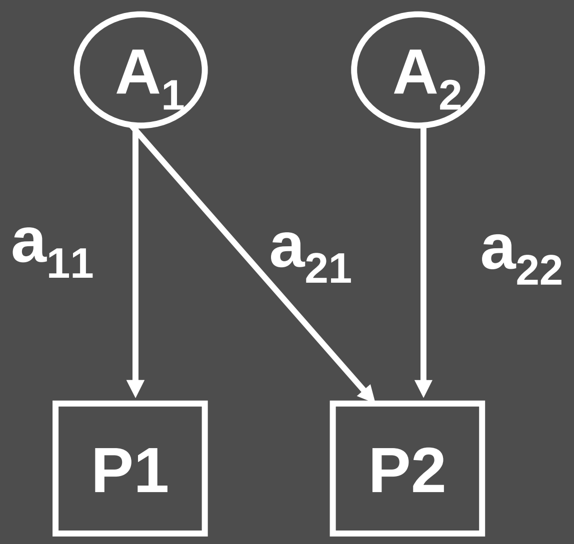
Cholesky Decomposition: Specification in Mx

Mx: Parameter Matrices

```
#define nvar 2
Begin Matrices;
X lower nvar nvar free          ! Genetic coefficients
Y lower nvar nvar free          ! C coefficients
Z lower nvar nvar free          ! E coefficients
G Full 1 nvar  free             ! means
End Matrices;
Begin Algebra;
A=X*X';                         ! Gen var/cov
C=Y*Y';                         ! C var/cov
E=Z*Z';                         ! E var/cov
P=A+C+E
End Algebra;
```

Within-Twin Covariances

Path Tracing



$$\Sigma_A = \begin{bmatrix} a^2_{11} & a_{11}a_{21} \\ a_{21}a_{11} & a^2_{21} + a^2_{22} \end{bmatrix}$$

X LOWER 2 × 2 ➡

$$\begin{array}{c|cc} & \mathbf{A1} & \mathbf{A2} \\ \hline \mathbf{P1} & a_{11} & 0 \\ \mathbf{P2} & a_{21} & a_{22} \end{array}$$

* or 'Star' Matrix Multiplication

$$\Sigma_A = X * X' = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} * \begin{bmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{bmatrix} = \begin{bmatrix} a^2_{11} + 0 \times 0 & a_{11}a_{21} + 0 \times a_{22} \\ a_{21}a_{11} + 0 \times a_{22} & a^2_{21} + a^2_{22} \end{bmatrix}$$

$$\Sigma_A = X * X' = \begin{bmatrix} a^2_{11} & a_{11}a_{21} \\ a_{21}a_{11} & a^2_{21} + a^2_{22} \end{bmatrix} \quad \Sigma_C = Y * Y' = \begin{bmatrix} c^2_{11} & c_{11}c_{21} \\ c_{21}c_{11} & c^2_{21} + c^2_{22} \end{bmatrix}$$

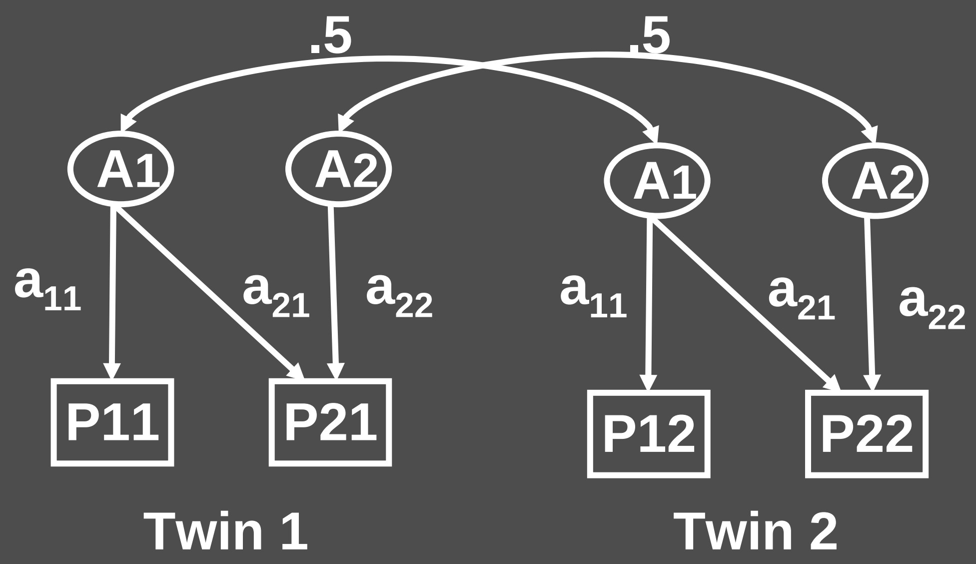
$$\Sigma_E = Z * Z' = \begin{bmatrix} e^2_{11} & e_{11}e_{21} \\ e_{21}e_{11} & e^2_{21} + e^2_{22} \end{bmatrix}$$

$$\Sigma_P = \Sigma_A + \Sigma_C + \Sigma_E$$

By rule of matrix addition:

$$\Sigma_P = \begin{bmatrix} a^2_{11} + c^2_{11} + e^2_{11} & a_{11}a_{21} + c_{11}c_{21} + e_{11}e_{21} \\ a_{21}a_{11} + c_{21}c_{11} + e_{11}e_{21} & a^2_{21} + a^2_{22} + c^2_{21} + c^2_{22} + e^2_{21} + e^2_{22} \end{bmatrix}$$

Cross-Twins Covariances, Genetic effects (DZ)



Path Tracing

Within-Traits (diagonals):

P_{11}-P_{12}= $.5\ a_{11}^2$

P_{21}-P_{22}= $.5\ a_{22}^2 + .5\ a_{21}^2$

Cross-Traits:

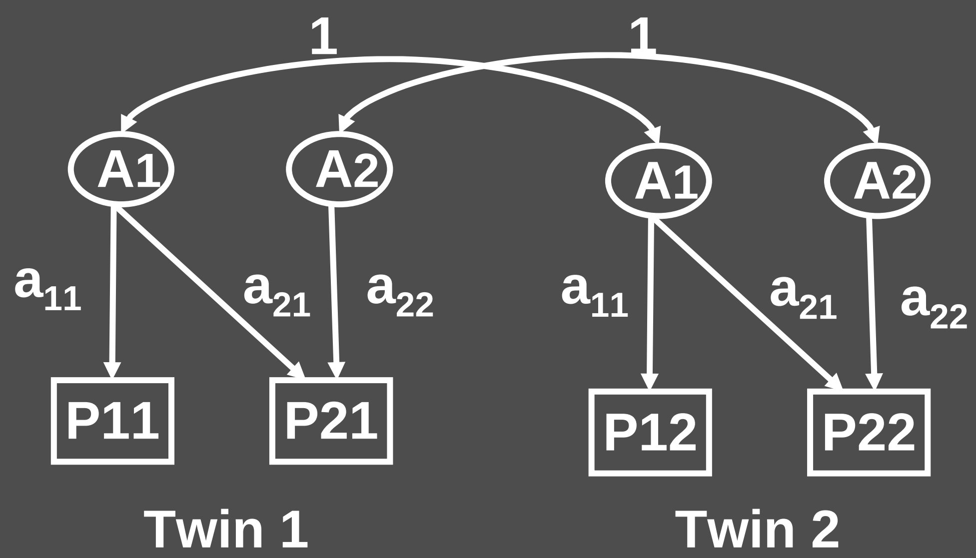
P_{11}-P_{22} = $.5\ a_{11}\ a_{21}$

P_{21}-P_{12} = $.5\ a_{21}\ a_{11}$

Kronecker Product ⊗ → @

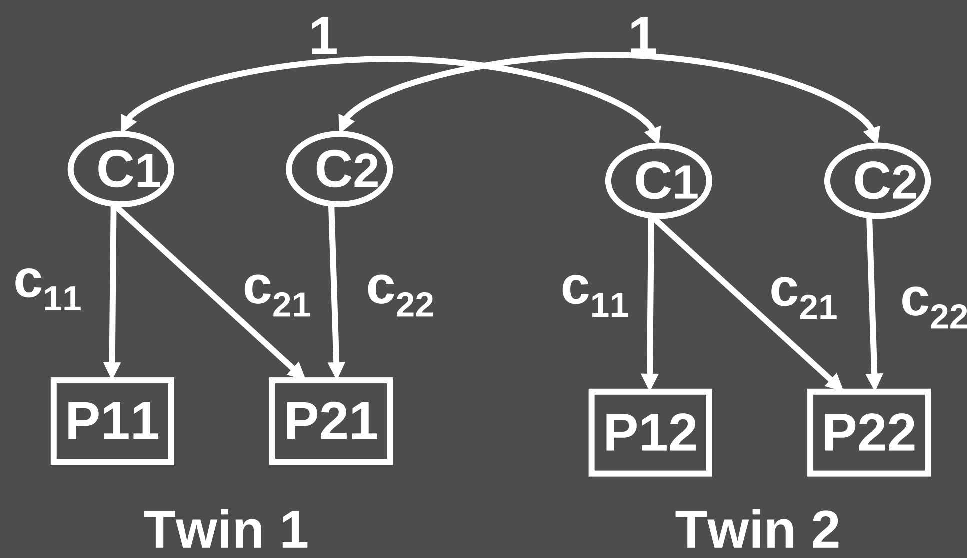
$$.5 \otimes \Sigma_A = .5 \otimes X * X' = \begin{bmatrix} .5\ a^2{}_{11} & .5\ a_{11}a_{21} \\ .5\ a_{21}a_{11} & .5\left(a^2{}_{21} + a^2{}_{22}\right) \end{bmatrix}$$

Cross-Twins Covariances, Genetic effects (MZ)



$$1 \otimes \Sigma_A = 1 \otimes X * X' = \begin{bmatrix} a^2_{11} & a_{11}a_{21} \\ a_{21}a_{11} & a^2_{21} + a^2_{22} \end{bmatrix}$$

Cross-Twins Covariances, C effects (MZ and DZ)



$$1 \otimes \Sigma_C = 1 \otimes Y * Y' = \begin{bmatrix} c^2_{11} & c_{11}c_{21} \\ c_{21}c_{11} & c^2_{21} + c^2_{22} \end{bmatrix}$$

Covariance Model for Twin pairs

MZ

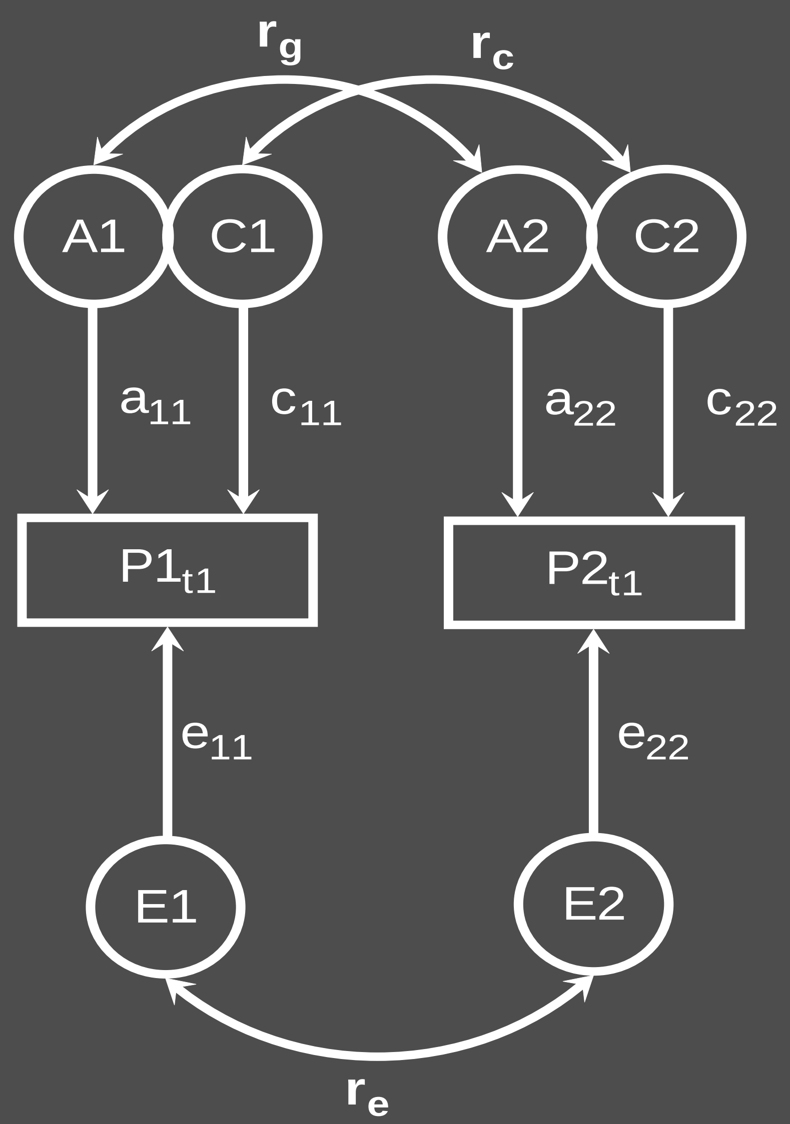
Covariance	A+C+E	| A+C _
	A+C	| A+C+E /

DZ

Covariance	A+C+E	| H@A+C _
	H@A+C	| A+C+E /

Standardized Estimates

Correlated Factors Solution



Covariances to Correlations

$$r_{12} = \frac{\sigma^2_{12}}{\sqrt{\sigma^2_{11} * \sigma^2_{22}}} \quad \Rightarrow \quad r_{12} = \frac{1}{\sqrt{\sigma^2_{11}}} * \sigma^2_{12} * \frac{1}{\sqrt{\sigma^2_{22}}}$$

In matrix form:

$$\begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{11}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{22}}} \end{bmatrix} * \begin{bmatrix} \sigma^2_{11} & \sigma^2_{12} \\ \sigma^2_{21} & \sigma^2_{22} \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{11}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{22}}} \end{bmatrix}$$

Genetic Correlations

$$\Sigma_A = \begin{bmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & r_G \\ r_G & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{A11}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{A22}}} \end{bmatrix} * \begin{bmatrix} \sigma^2_{A11} & \sigma^2_{A12} \\ \sigma^2_{A21} & \sigma^2_{A22} \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{A11}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{A22}}} \end{bmatrix}$$

Specification in Mx

$$\begin{bmatrix} 1 & r_G \\ r_G & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{A11}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{A22}}} \end{bmatrix} * \begin{bmatrix} \sigma^2_{A11} & \sigma^2_{A12} \\ \sigma^2_{A21} & \sigma^2_{A22} \end{bmatrix} * \begin{bmatrix} \frac{1}{\sqrt{\sigma^2_{A11}}} & 0 \\ 0 & \frac{1}{\sqrt{\sigma^2_{A22}}} \end{bmatrix}$$

1 **Matrix Function in Mx: R = \stnd (A);**

2 **R = \sqrt(I . A)~ * A * \sqrt(I . A)~;**

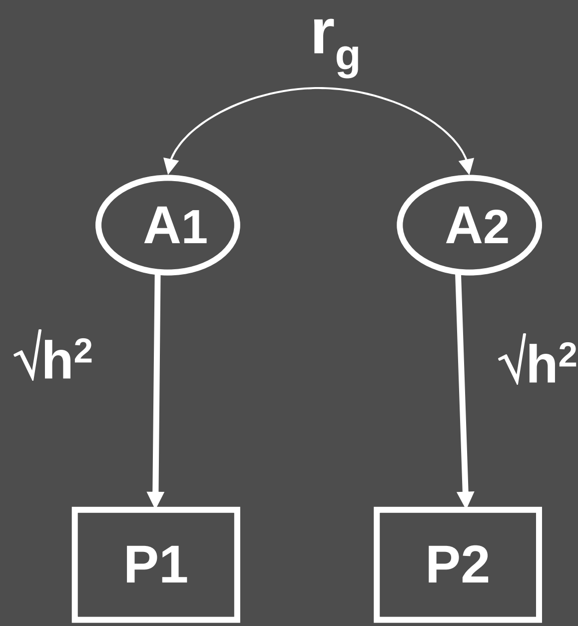
Where I is an Identity matrix : $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ **And I.A =** $\begin{pmatrix} \sigma^2_{A11} & 0 \\ 0 & \sigma^2_{A22} \end{pmatrix}$

Interpretation

- **High Genetic correlation: large overlap in genetic effects on the two traits**

- **Does it mean that the phenotypic correlation between the traits is largely due to genetic effects?**

    - **No,** the substantive importance of a particular r_g depends both on the value of the correlation and the value of the A paths, i.e. the heritabilities of both traits.



$$(\sqrt{h^2}_{p1} * r_g * \sqrt{h^2}_{p2}) / r_{PH}$$

$$(\sqrt{.63} * -0.525 * \sqrt{.33}) / -0.29 = .8357$$

Interpretation

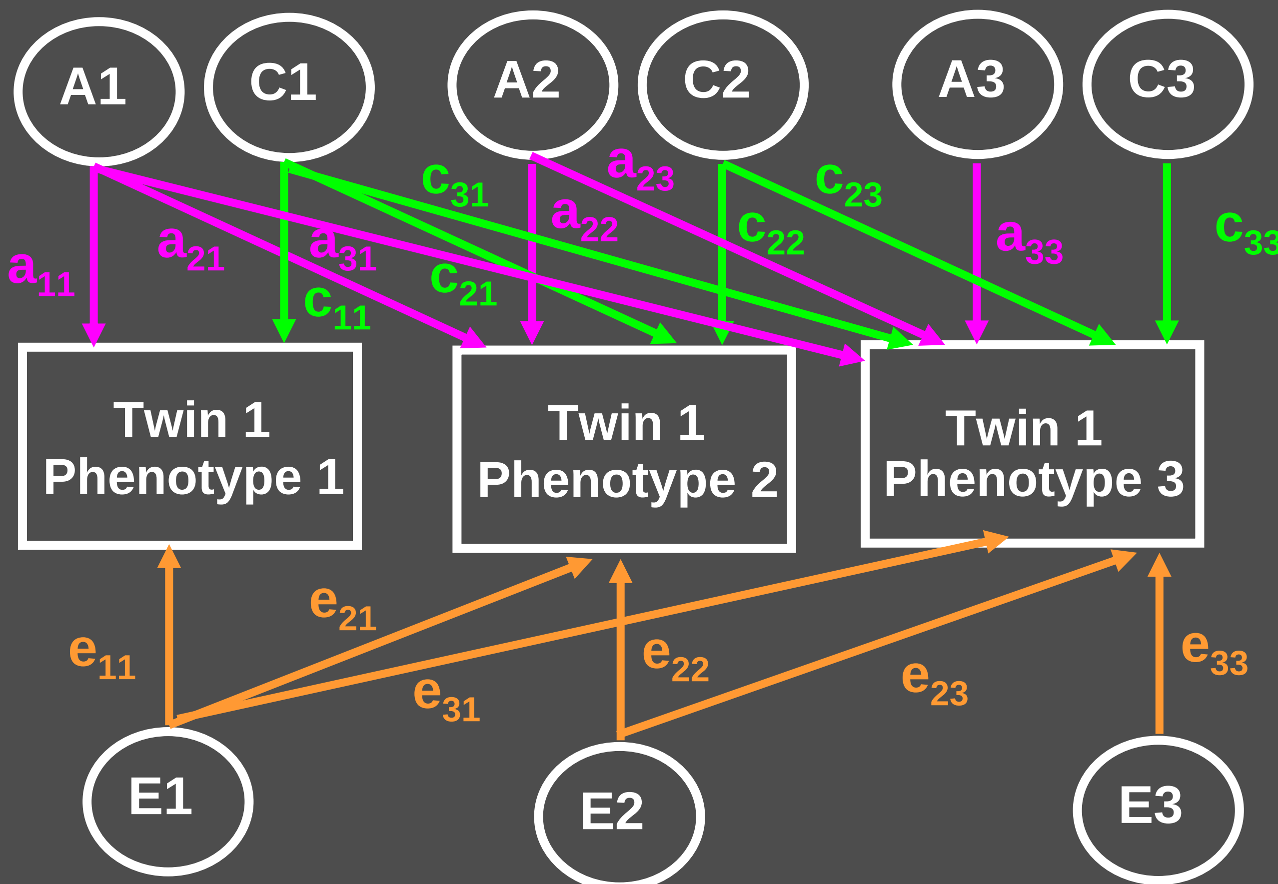
- For Example, consider a phenotypic correlation of 0.40

With $\sqrt{h^2_{p1}} = .7$ and $\sqrt{h^2_{p2}} = .6$ with rg = .3
.19 (49%) of the phenotypic correlation can be attributed to additive genetic effects.

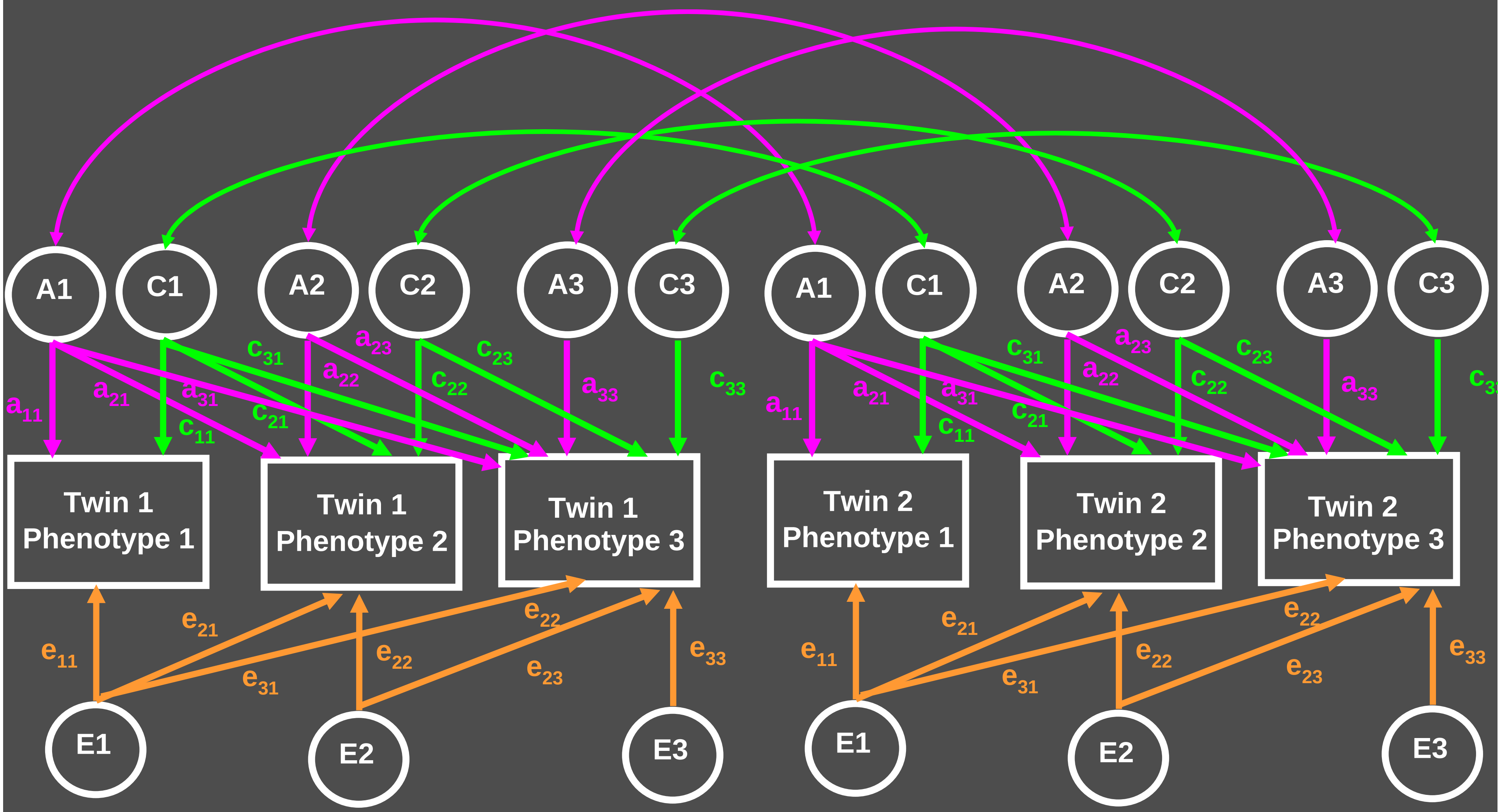
With $\sqrt{h^2_{p1}} = .2$ and $\sqrt{h^2_{p2}} = .4$ with rg = .8
.20 (49%) of the phenotypic correlation can be attributed to additive genetic effects.

- Weakly heritable traits can still have a large portion of their correlation attributable to genetic effects.

More variables……



More variables……



Mx: Parameter Matrices

```
#define nvar 3
Begin Matrices;
X lower nvar nvar free          ! Genetic coefficients
Y lower nvar nvar free          ! C coefficients
Z lower nvar nvar free          ! E coefficients
G Full 1 nvar  free             ! means
End Matrices;
Begin Algebra;
A=X*X';                         ! Gen var/cov
C=Y*Y';                         ! C var/cov
E=Z*Z';                         ! E var/cov
P=A+C+E
End Algebra;
```

X LOWER 3 × 3 →

A1	A2	A3
a_{11}	0	0
a_{21}	a_{22}	0
a_{31}	a_{32}	a_{33}

Y LOWER 3 × 3 →

C1	C2	C3
c_{11}	0	0
c_{21}	c_{22}	0
c_{31}	c_{32}	c_{33}

Z LOWER 3 × 3 →

E1	E2	E3
e_{11}	0	0
e_{21}	e_{22}	0
e_{31}	e_{32}	e_{33}