GWAS model extensions

Christian Werner

(Quantitative geneticist and biostatistician) **EiB, CIMMYT**, Texcoco (Mexico)

Filippo Biscarini

HerrFalloppio

(Biostatistician, bioinformatician and quantitative geneticist) **CNR-IBBA**, Milan (Italy)

Oscar González-Recio

OscarGenomics

(Computational biologist and quantitative geneticist) **INIA-UPM**, Madrid (Spain)

GWAS for other gene actions

Additive coding

Genotype	101 coding	012 coding	Centered coding
BB	$-a_i$	0	$-2p_ia_i$
Bb	0	a_i	$(1-2p_i)a_i$
bb	a_i	$2a_{i<}$	$(2-2p_i)a_i$

EXAMPLE

aa Bb
AA bb
Aa bb
aa bb

101 coding: $\boldsymbol{Za} = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$

012 coding: $\boldsymbol{Za} = \begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$

centered coding: $\boldsymbol{Za} = \begin{pmatrix} -0.75 & 0.75 \\ 1.25 & -0.25 \\ 0.25 & -0.25 \\ -0.75 & -0.25 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$

p_1=0.375; p_2=0.125;

Dominance coding

Genotype	010 coding	012 coding (optional additive)	Centered coding
BB	0	0	$-2p_i^2 d_i$
Bb	d_i	a_i	$2p_i q_i d_i$
bb	0	$2a_i$	$2q_i d_i$

EXAMPLE

$aa \quad Bb$

$AA \quad bb$

$Aa \quad bb$

$aa \quad bb$

101 coding

$$Zd = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Centered coding

$$Zd = \begin{pmatrix} 0.78 & 0.22 \\ -0.28 & 1.53 \\ 0.47 & 1.53 \\ 0.78 & 1.53 \end{pmatrix}$$

p_1=0.375; p_2=0.125;

A#A Epistasis coding

Genotype 1	AA (0)	Aa (1)	aa (2)
BB (0)	00	01	02
Bb (1)	10	11	12
bb (2)	20	21	22

CAUTION
This codification does not necessary keep linearity (covariates is not recommended, use factors, or separate 3 gene interactions (x_1={00,01,02};x_2={10,11,12};x_3={20,21,22})

Centered coding is also possible with appropriate products

A#D Epistasis coding

Genotype 1	AA (0)	Aa (1)	aa (0)
BB (0)	00	01	00
Bb (1)	10	11	10
bb (2)	20	21	20

CAUTION
This codification does not necessary keep linearity (covariates is not recommended, use factors, or separate 3 gene interactions (x_1={00,01,00};x_2={10,11,10};x_3={20,21,20})

D#D Epistasis coding

Genotype 1	AA (0)	Aa (1)	aa (0)
BB (0)	00	01	00
Bb (1)	10	11	10
bb (0)	00	01	00

CAUTION
This codification does not necessary keep linearity (covariates is not recommended, use factors, or separate 3 gene interactions (x_1={00,01,00};x_2={10,11,10};x_3={00,01,00})

Other additive codifications (A#A#A, A#D#A, ...)

- Just follow the same rationale

Genotype x Environment coding

Genotype 1	Teatment (T1)	Control
BB (0)	T_0	C_0
Bb (1)	T_1	C_1
bb (2)	T_1	C_2

CAUTION
This codification does not necessary keep linearity (covariates is not recommended, use factors, or separate 2 covariates (x_1={T0,T1,T2};x_2={C0,C1,C2})

Centered coding is also possible with appropriate products

Genotype x Environment coding