Tareas Octubre 25

October 27, 2019
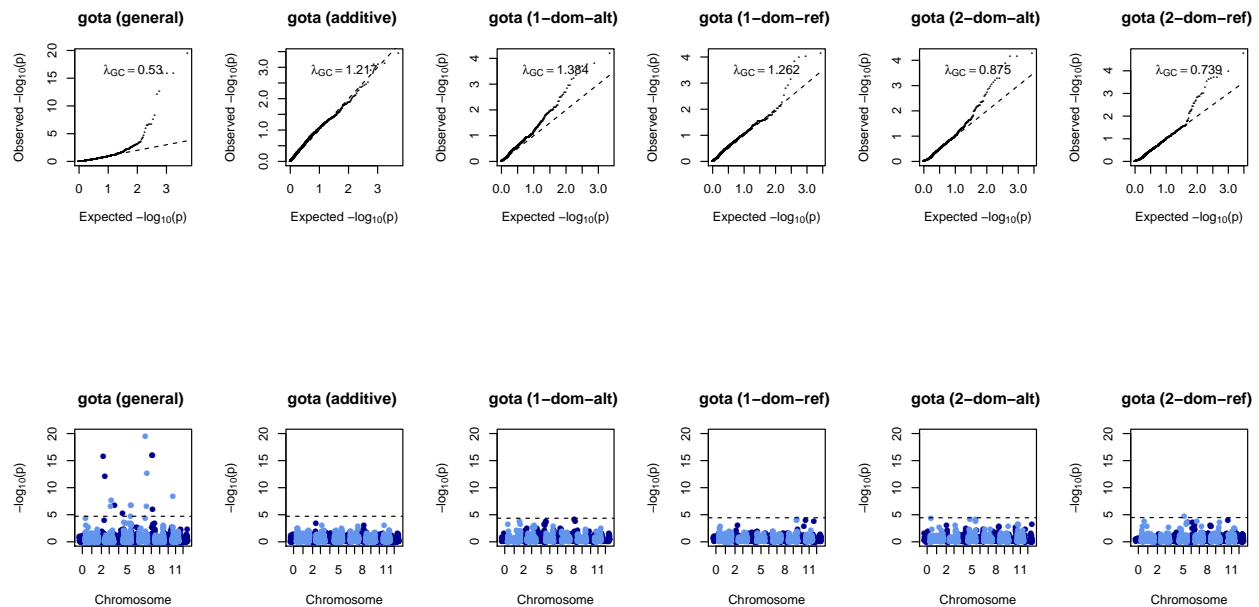
- Run GWAS with Enciso's data
- Genotype coding from tetraploid to diploid
- Write the goodness of fit for each model
- Read paper GWAS toolsλ
- Study GWAS models

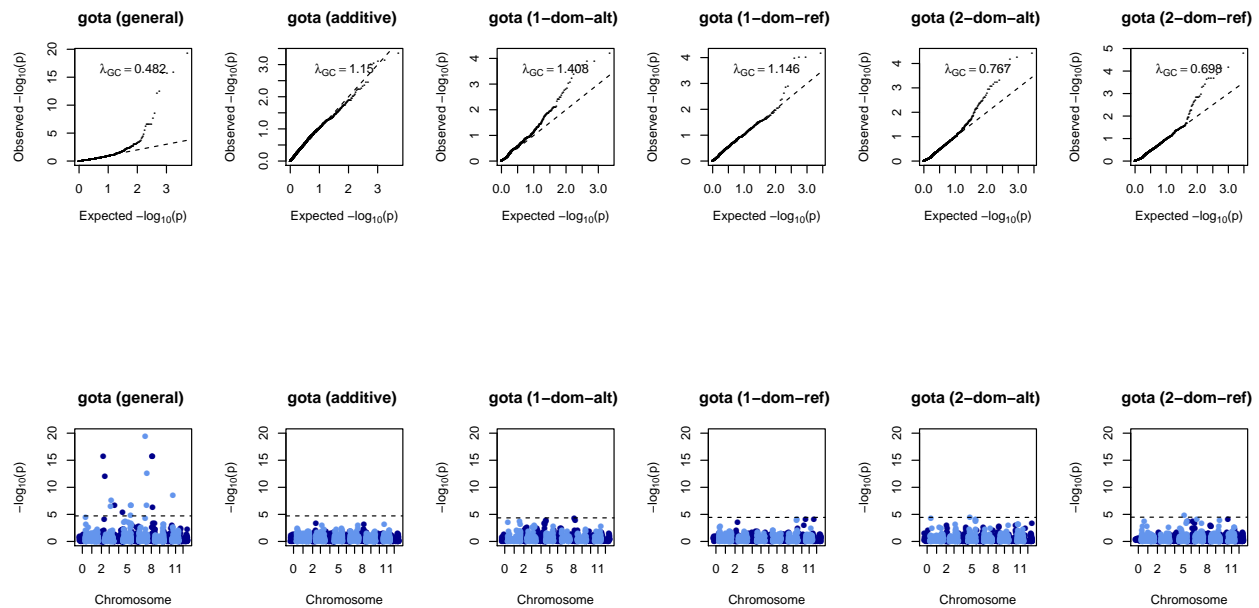## 1 Goodness of fit for each model

The most used way to show the goodness of fit for a gwas model is to compute the genomic inflation factor λ_{GC} which describes the deviation of the median of the observed test statistics from the median of the expected test statistics [Smith2009]. "Considering the assumption that most of the SNPs are not associated with the trait, λ_{GC} should be close to one if no confounder is present. If λ GC is inflated (larger than 1) or deflated (lower than 1), a correction for population structure is recommended to avoid false positives or false negatives, respectively". [Smith2009].
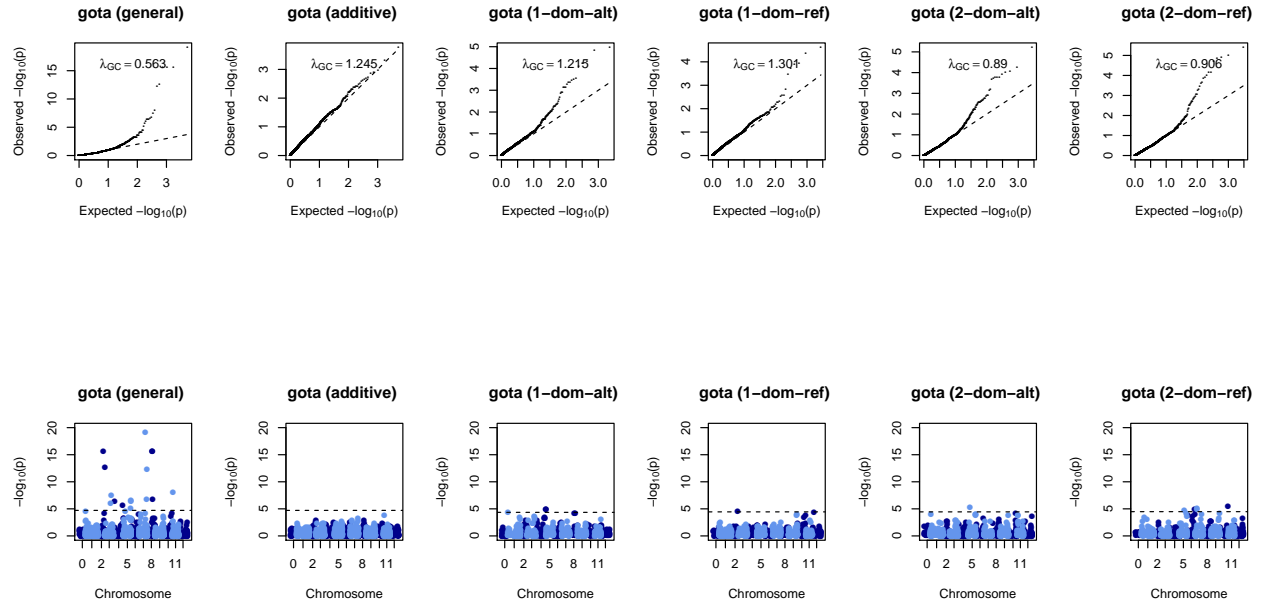
2 GWAS with Agrosavia data (tetraploid)
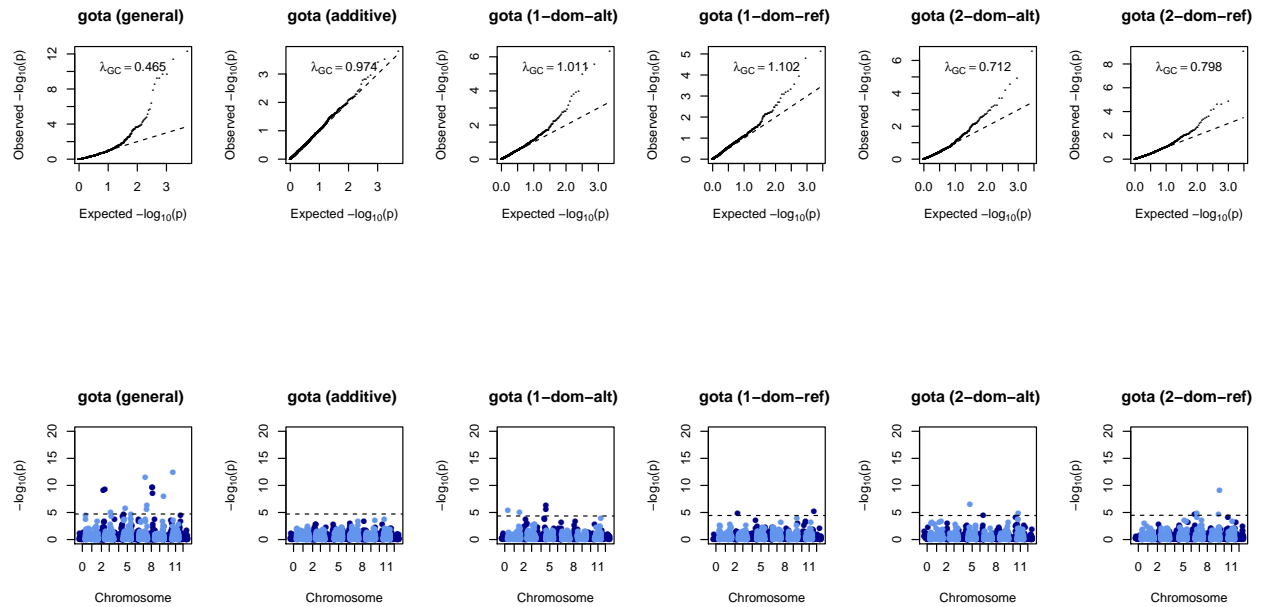
GWAS 4–ploidy with Naive for gota trait

GWAS 4–ploidy with Kinship for gota trait

GWAS 4−ploidy with Structure for gota trait

gota (general)
λ_GC = 0.563

gota (additive)
λ_GC = 1.245

gota (1−dom−alt)
λ_GC = 1.215

gota (1−dom−ref)
λ_GC = 1.301

gota (2−dom−alt)
λ_GC = 0.89

gota (2−dom−ref)
λ_GC = 0.906

gota (general)
gota (additive)
gota (1−dom−alt)
gota (1−dom−ref)
gota (2−dom−alt)
gota (2−dom−ref)

GWAS 4−ploidy with Kinship+Structure for gota trait

gota (general)
λ_GC = 0.465

gota (additive)
λ_GC = 0.974

gota (1−dom−alt)
λ_GC = 1.011

gota (1−dom−ref)
λ_GC = 1.102

gota (2−dom−alt)
λ_GC = 0.712

gota (2−dom−ref)
λ_GC = 0.798

gota (general)
gota (additive)
gota (1−dom−alt)
gota (1−dom−ref)
gota (2−dom−alt)
gota (2−dom−ref)

Model	Thrsh	Marker	Chr	Position	Ref	Alt	Score	Effect	Annotation	Best.Arabidopsis
general	4.72	c1_12386	2	37411847	0	1	9.18	NA	Zinc finger protein	AT1G30970.1
general	4.72	c1_5895	2	48564909	0	1	9.27	NA	Acetyl-CoA synthetase	AT5G36880.2
general	4.72	c1_6869	3	41992129	0	1	5.1	NA	CDPK	AT1G12580.1
general	4.72	c2_32854	5	10418984	0	1	5.75	NA	Signal transducer	AT1G67900.1
general	4.72	c2_3512	5	51477749	0	1	4.75	NA	Conserved gene of unknown function	AT3G04350.1
general	4.72	c2_15929	7	43192596	0	1	11.41	NA	Conserved gene of unknown function	AT2G04280.1
general	4.72	c2_12404	7	52544646	0	1	5.57	NA	Conserved gene of unknown function	AT1G78810.1
general	4.72	c2_12601	7	53106007	0	1	6.37	NA	3-hydroxybutyryl-CoA dehydrogenase	AT3G15290.1
general	4.72	c2_49245	8	35651378	0	1	9.64	NA	UDP-sulfoquinovose synthase	AT4G33030.1
general	4.72	c2_41463	8	35922019	0	1	9.64	NA	Conserved gene of unknown function	AT2G26110.1
general	4.72	c1_821	8	37494529	0	1	8.51	NA	Conserved gene of unknown function	AT4G32820.1
general	4.72	c2_2998	9	58689196	0	1	7.91	NA	Gene of unknown function	No_Hit
general	4.72	c1_11246	11	4745919	0	1	12.35	NA	Seed maturation protein PM36	AT3G16990.1
1-dom-alt	4.35	c2_36664	1	535454	0	1	5.35	-0.59	Adagio protein 3	AT1G68050.1
1-dom-alt	4.35	c2_4875	1	83177518	0	1	4.98	-0.59	Conserved gene of unknown function	AT5G65810.1
1-dom-alt	4.35	c1_3484	4	71945107	0	1	6.35	-0.78	Cinnamoyl-CoA reductase	AT2G33590.1
1-dom-alt	4.35	c2_10568	4	71954528	0	1	5.55	-0.69	3-hydroxy-3-methylglutaryl coenzyme A reductase	AT1G76490.1
1-dom-ref	4.44	c1_7325	2	42764457	0	1	4.79	0.64	POM30	AT5G67500.1
1-dom-ref	4.44	c2_40748	12	12654160	0	1	5.14	0.57	Pentatricopeptide repeat-containing protein	AT2G37230.1
2-dom-alt	4.44	c2_32854	5	10418984	0	1	6.58	0.72	Signal transducer	AT1G67900.1
2-dom-alt	4.44	c1_6992	6	53677174	0	1	4.55	-0.45	Conserved gene of unknown function	AT1G08760.1
2-dom-alt	4.44	c1_15787	11	6943511	0	1	4.91	-0.55	Pectinesterase 51	AT5G09760.1
2-dom-ref	4.48	c2_24064	6	54713088	0	1	4.65	-0.44	Bell-like homeodomain protein 2	No_Hit
2-dom-ref	4.48	c1_10855	7	9998099	0	1	4.84	0.55	DNA repair protein RAD51 homolog	AT5G20850.1
2-dom-ref	4.48	c1_11907	9	51386555	0	1	4.61	0.42	ATP binding protein	AT5G35960.1
2-dom-ref	4.48	c2_2998	9	58689196	0	1	9.1	-0.76	Gene of unknown function	No_Hit

GWAS 2−ploidy with Kinship+Structure for gota trait

3 GWAS with Agrosavia data (tetraploid)

GWAS 2−ploidy with Naive for gota trait

GWAS 2−ploidy with Kinship for gota trait

GWAS 2–ploidy with Structure for gota trait
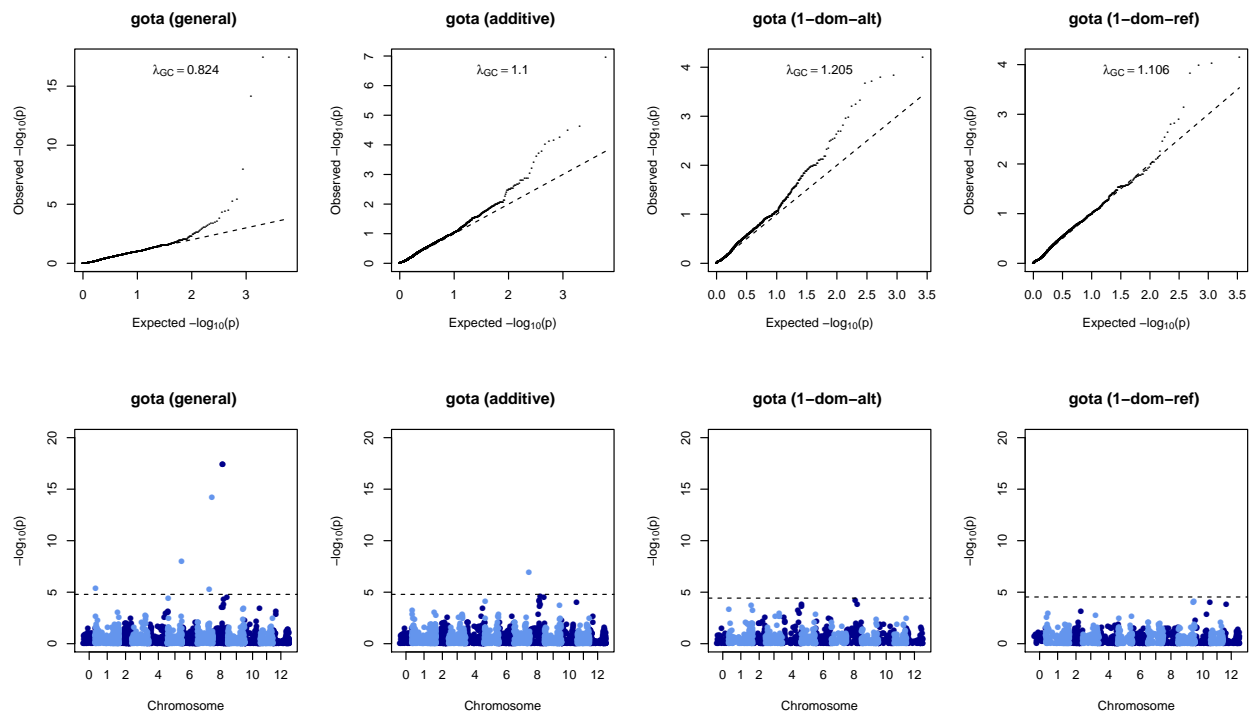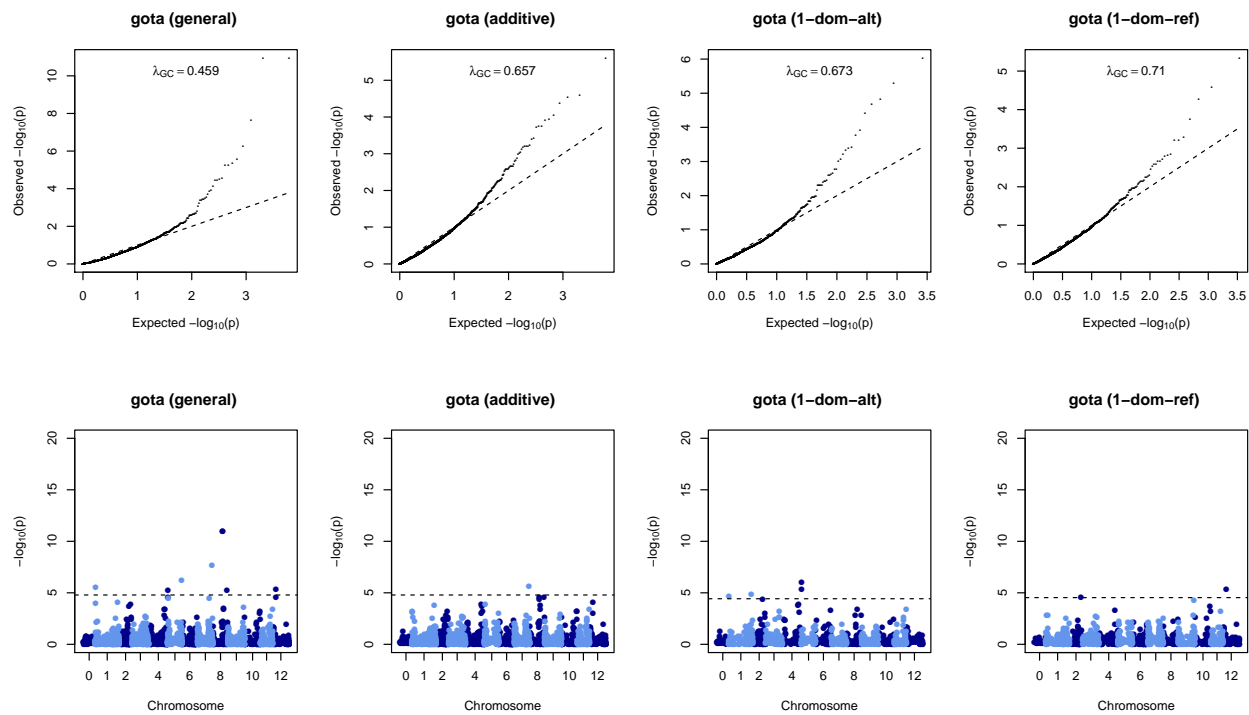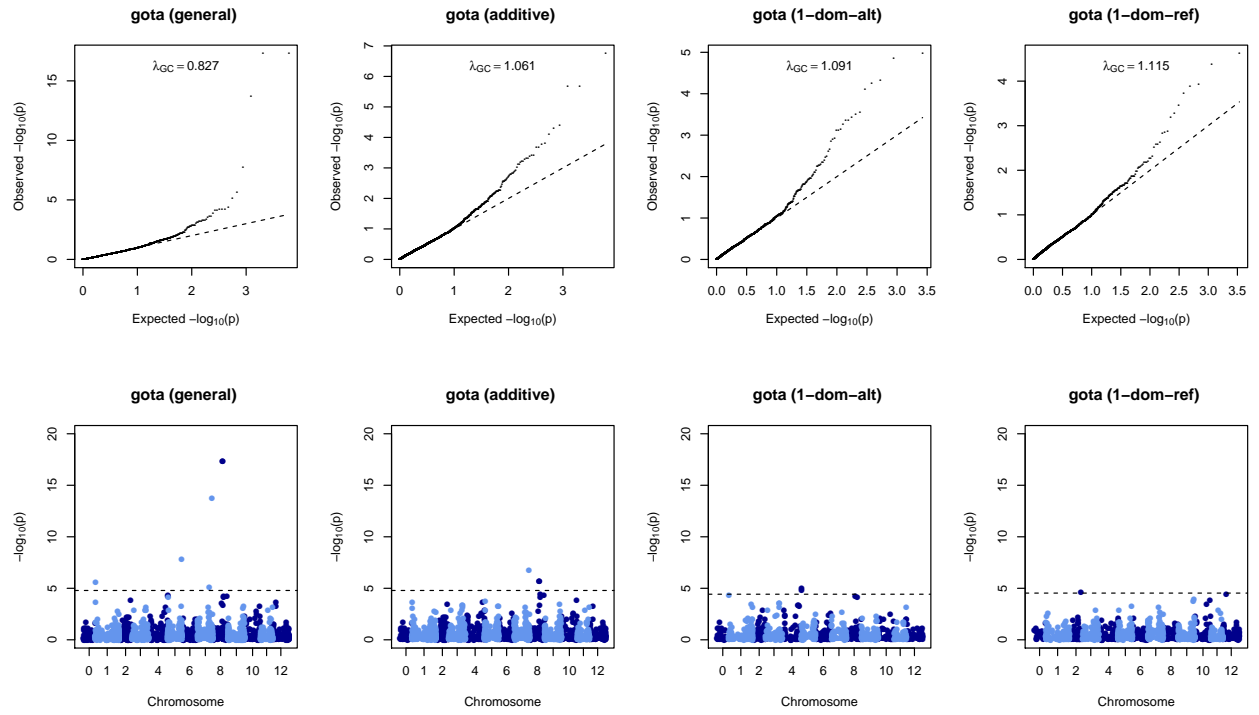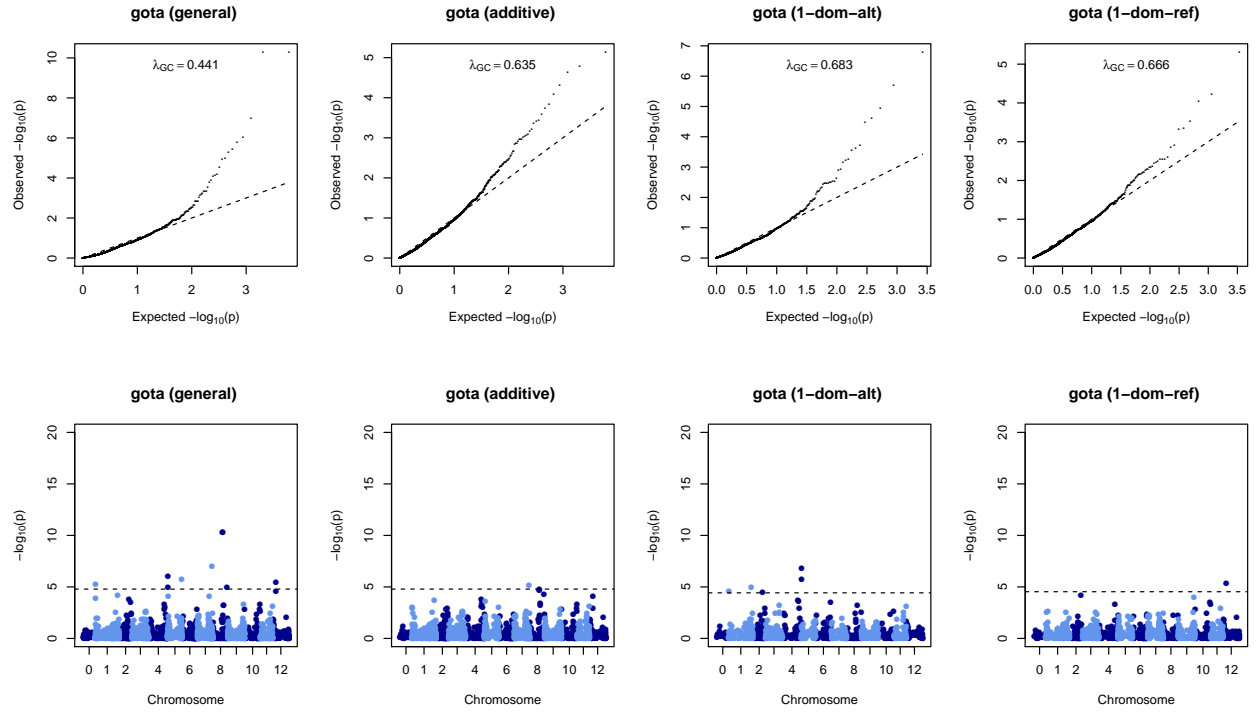

GWAS 2–ploidy with Kinship+Structure for gota trait

Model	Thresh	Marker	Chr	Position	Ref	Alt	Score	Effect	Annotation	Best.Arabidopsis
general	4.79	c2_36660	1	534626	0	1	5.31	NA	Circadian clock-associated FKF1	AT1G68050.1
general	4.79	c1_3484	4	71945107	0	1	6.02	NA	Cinnamoyl-CoA reductase	AT2G33590.1
general	4.79	c2_10568	4	71954528	0	1	4.94	NA	3-hydroxy-3-methylglutaryl coenzyme A reductase	AT1G76490.1
general	4.79	c2_3512	5	51477749	0	1	5.79	NA	Conserved gene of unknown function	AT3G04350.1
general	4.79	c2_12601	7	53106007	0	1	6.97	NA	3-hydroxybutyryl-CoA dehydrogenase	AT3G15290.1
general	4.79	c2_49245	8	35651378	0	1	10.29	NA	UDP-sulfoquinovose synthase	AT4G33030.1
general	4.79	c2_41463	8	35922019	0	1	10.3	NA	Conserved gene of unknown function	AT2G26110.1
general	4.79	c2_19081	8	52490785	0	1	4.98	NA	L-lactate dehydrogenase	AT4G17260.1
general	4.79	c2_40745	12	12655349	0	1	5.42	NA	Pentatricopeptide repeat-containing protein	AT2G37230.1
additive	4.79	c2_12601	7	53106007	0	1	5.14	-0.47	3-hydroxybutyryl-CoA dehydrogenase	AT3G15290.1
additive	4.79	c2_49245	8	35651378	0	1	4.79	0.54	UDP-sulfoquinovose synthase	AT4G33030.1
1-dom-alt	4.42	c2_36664	1	535454	0	1	4.61	-0.56	Adagio protein 3	AT1G68050.1
1-dom-alt	4.42	c2_4875	1	83177518	0	1	4.93	-0.63	Conserved gene of unknown function	AT5G65810.1
1-dom-alt	4.42	c2_49495	2	37017928	0	1	4.48	-0.52	Protease C56	AT3G02720.1
1-dom-alt	4.42	c1_3484	4	71945107	0	1	6.8	-0.9	Cinnamoyl-CoA reductase	AT2G33590.1
1-dom-alt	4.42	c2_10568	4	71954528	0	1	5.72	-0.76	3-hydroxy-3-methylglutaryl coenzyme A reductase	AT1G76490.1
1-dom-ref	4.53	c2_40748	12	12654160	0	1	5.31	0.6	Pentatricopeptide repeat-containing protein	AT2G37230.1