BMC Proceedings

BioMed Central

Proceedings

Open Access

A comparative study of three methods for detecting association of quantitative traits in samples of related subjects

Aude Saint Pierre*, Zulma Vitezica and Maria Martinez

Address: INSERM, U.563, University Paul-Sabatier, CPTP, Toulouse F-31300, France

E-mail: Aude Saint Pierre* - Aude.saint-pierre@inserm.fr; Zulma Vitezica - Zulma.vitezica@inserm.fr; Maria Martinez - Maria.martinez@inserm.fr
*Corresponding author

from Genetic Analysis Workshop 16
St Louis, MO, USA 17-20 September 2009

Association Studies in Families

Tests for genetic association can use family data when relatedness between individuals is modeled appropriately (e.g. George and Elston, 1987)

Linear Mixed (polygenic) Model for Association

Vector of quantitative phenotype *Y*

$$Y = \mu + \mathbf{bg} + \mathbf{G} + \mathbf{e}$$

g: genotype indicator vector gi in {0,1,2}

b: (b_b, b_w) additive fixed effect of the allele

- *Similarities between individuals*

G: random polygenic effect ~ MVN(0, $\boldsymbol{\Phi\sigma}_G^2$)

e: random residual effect ~ MVN(0, $\mathbf{I}\sigma_e^2$)

- Maximum Likelihood (ML)
- Software packages (**SOLAR**,MERLIN, …)

Controlling for Stratification

- If stratum were known…
  - For each individual genotype (g_{ij})
  - Average number of alleles in a strata (b_{ij})
  - Adjust for stratum differences ($w_{ij} = g_{ij} - b_{ij}$)

$$Y_{ij} = \mu + \hat{\beta}_b b_{ji} + \hat{\beta}_w w_{ij}$$

- How to define stratum then?
  Use family data to estimate b_{ij}

Extended Families (Fi)

$$b_{ij} = \begin{cases} \dfrac{b_{iF_j} + b_{iM_j}}{2} & \text{average of parental controls} \\ \displaystyle\sum_{k}^{sibship} \frac{g_{ik}}{n_{sibs}} & \text{average of sibling genotypes} \\ g_{ij} & \text{self - genotype} \\ undefined & \text{otherwise} \end{cases}$$

Non-founders, $w_{ij} = g_{ij} - b_{ij}$

Tests of association

Mixed model (μ , b_b, b_w, σ^2_G, σ^2_e)

Orthogonal components

		$\underline{b_b}$	$\underline{b_w}$	
MG	L0	[0]	[0]	
	L1	x	[=b_b]	Test of Stratification
QTDT	L2	x	[0]	
	L3	x	x	

LRT ~ χ^2 1df

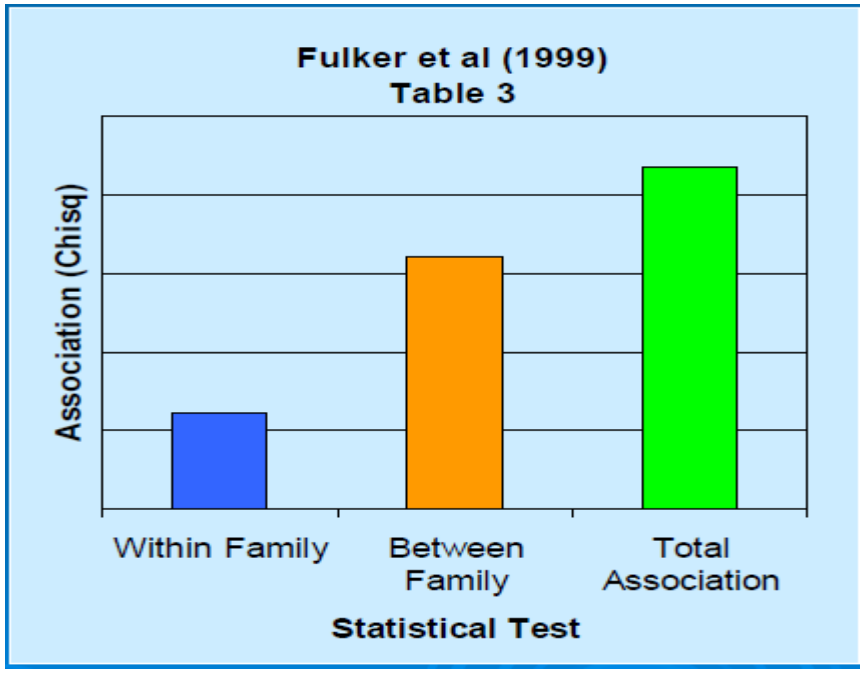
Transmission Disequilibrium Test -- Effective sample size

- **Discards data** on the relatives not fulfilling either one of the 2 conditions

(1) Both parents genotyped and at least one of them is heterozygote

(2) They have at least one sibling with a different genotype

- **Focus on** within family component of association



To Compare Methods for association analysis for quantitative traits in related individuals

- Type I error & Power:
  - ***Quantitative Transmission-Desequilibrium Test*** (QTDT)
  - *Quantitative Trait Linkage Disequilibrium Test* (QTLD)
  - ***Measured Genotype*** (MG)

➢ These approaches differ in the amount and type of marker information used for testing association.

Pedigree Data – Framingham Heart Study

704 pedigrees with ≥ 2 non-founders individuals with available phenotype & genotype data
(out of 12,407 subjects: 6,009 have phenotype data; ~48% have genotype data)

SNP Data: Affy GeneChip Human Mapping 500K Array.

QC steps:

(1) Exclude SNPs: call rate <95%; monomorphic or low MAF (<1%) ; significant (p-value<10^{-6}) departure to Hardy-Weinberg equilibrium (using unrelated subjects); Mendelian consistency checked with Pedstat [Wigginton & Abecasis 2005]

(2) DNA samples with <95% call rate: all genotypes zeroed out.

Simulated traits

Heritability (%)

HDL: 54

TG: 38

HDL: 0.2% (ADD)

HDL: 0.3% (DOM)

TG: 0.4% (ADD)

HDL: 0.3% (ADD)

TG: 0.3% (ADD)

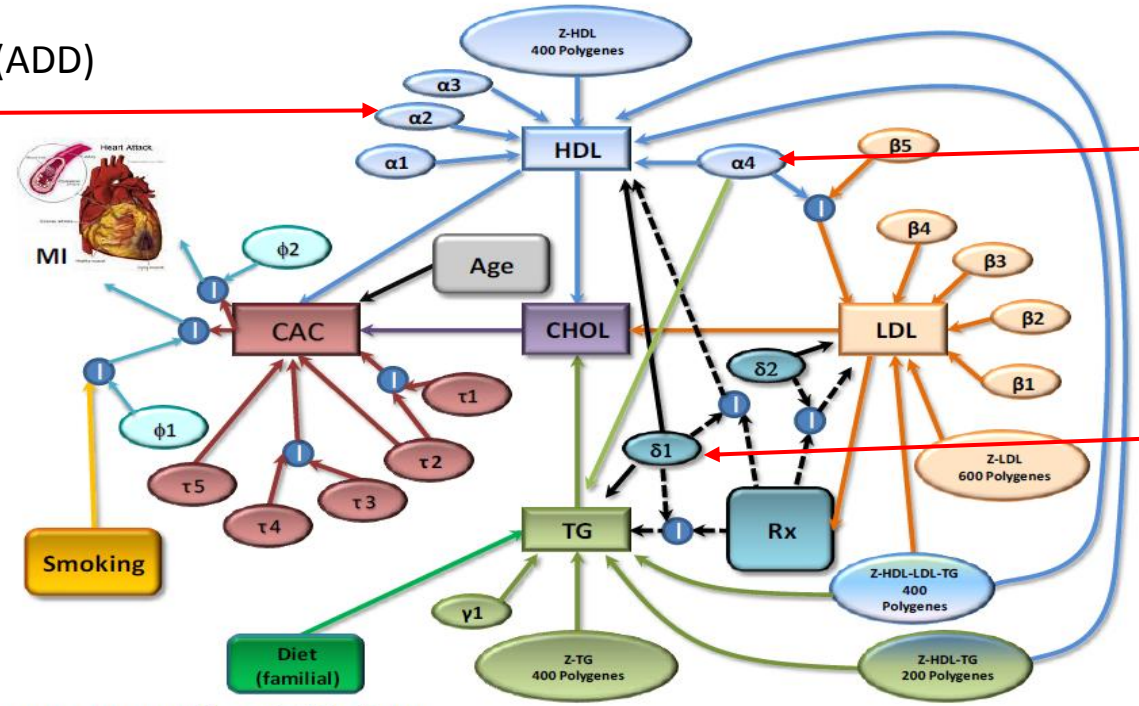


Figure 1. GAW16 Problem 3 simulation diagram.

200 replicates of FHS pedigree sample

Characteristics of the tested (causative and non-causative) SNPs

						h^2g		No Ind. with genotype
Chr	**Gene**	**Pos (bp)**	**SNP**	**MAF %**	**D' (with causal)***	**HDL**	**TG**	
7	None	24 734 008	rs2521760	12.7	-	-	-	5826
8	**alpha4**	19 794 163	rs17091651	10.0	0.04 (alpha4)	-	-	5945
		19 868 351	**rs3200218**	**21.7**		**0.3%**	**0.4%**	5854
		19 943 326	rs4244457	32.9	0.04 (alpha4)	-	-	5962
19	**alpha2**	46 010 146	rs11083567	18.2	0.07 (alpha2) - 0.03 (delta1)	-	-	5951
		46 089 501	**rs8103444**	**24.4**	0.003 (delta1)	**0.2%**	-	5995
	delta1	46 210 613	**rs8192719**	**24.9**	0.003 (alpha2)	**0.3%**	**0.3%**	5995
		46 335 684	rs1631931	13.5	0.01 (alpha2) - 0.03 (delta1)	-	-	5990

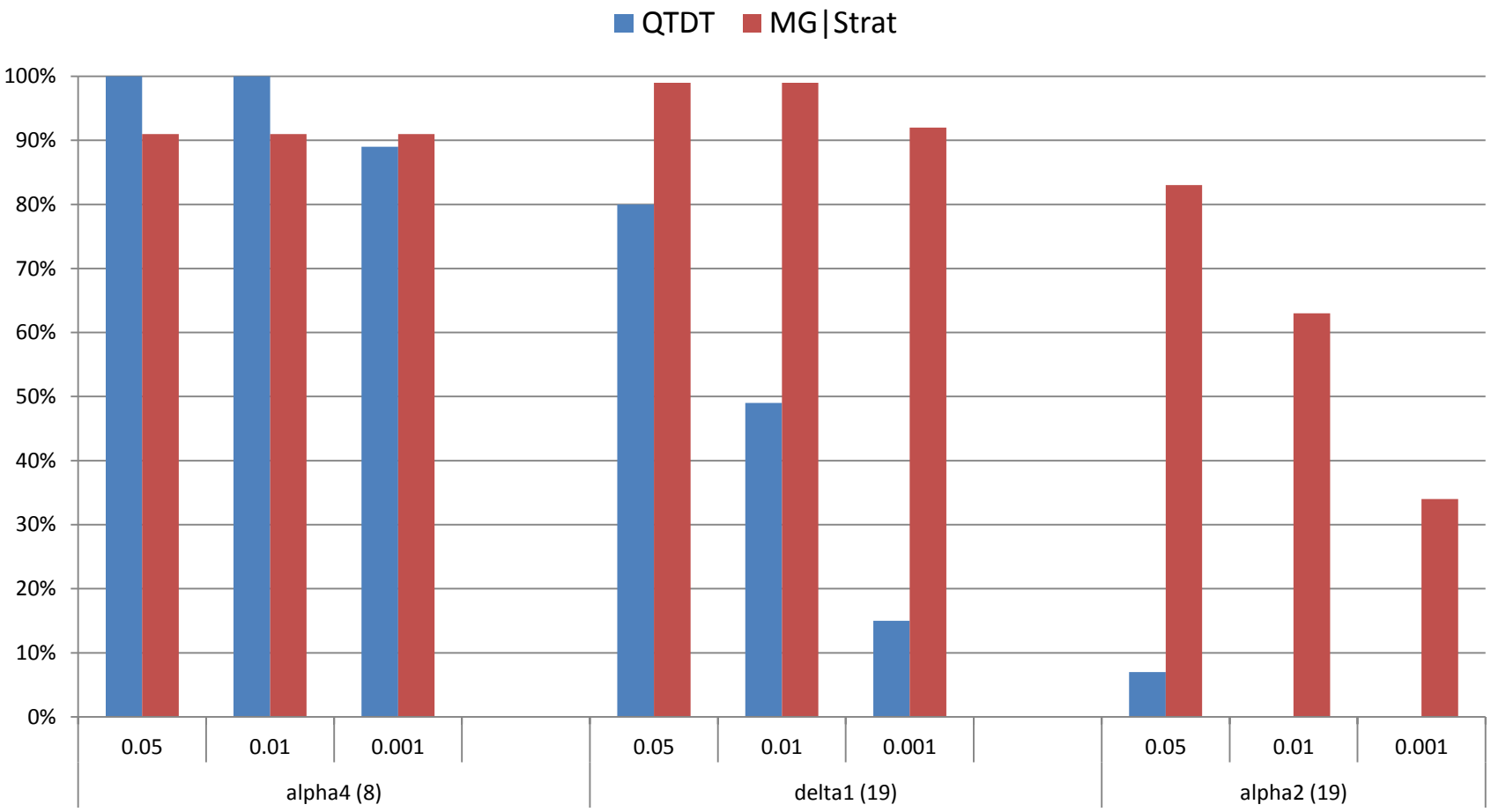
* pairwise linkage disequilibrium coefficient (D/Dmax) between the functional variant (symbol) and the SNPs in its vicinity (<200kb)

Type I error –
Mean χ^2 statistics (μ-χ^2) and rate of significant Association results

				μ-χ^2 (sd)			P=5%	
Chr	**Gene**	**SNP**	**Trait**	**QTDT**	**MG**	**MG_S**	**QTDT**	**MG_S**
				A- No Linkage & No Association				
7	none	rs2521760	HDL	0.48 (0.62)	0.73 (0.85)	0.72 (0.86)	0%	0%
			TG	0.86 (1.10)	0.64 (0.70)	0.60 (0.69)	3%	1%
			TG_Diet	0.87 (1.00)	0.62 (0.64)	0.56 (0.63)	4%	0%
			TG_Rob	0.99 (1.28)	0.48 (0.63)	0.44 (0.62)	4%	0%
				B- Linkage & No association				
8	alpha4	rs4244457	HDL	1.50 (1.79)	0.63 (0.79)	0.51 (0.75)	6%	1%
			TG	1.76 (1.86)	0.65 (1.13)	0.61 (1.12)	**14%**	3%
			TG_Rob	1.52 (1.74)	0.52 (0.98)	0.49 (0.99)	9%	3%
19	alpha2 /delta1	rs11083567	HDL	0.85 (0.97)	0.40 (0.51)	0.37 (0.52)	2%	0%
19	delta1	rs1631931	HDL	1.86 (2.02)	1.08 (1.24)	0.99 (1.19)	**12%**	1%
			TG	0.93 (1.35)	0.62 (0.79)	0.61 (0.79)	6%	1%
			TG_Rob	0.97 (1.29)	0.75 (0.93)	0.74 (0.93)	2%	1%

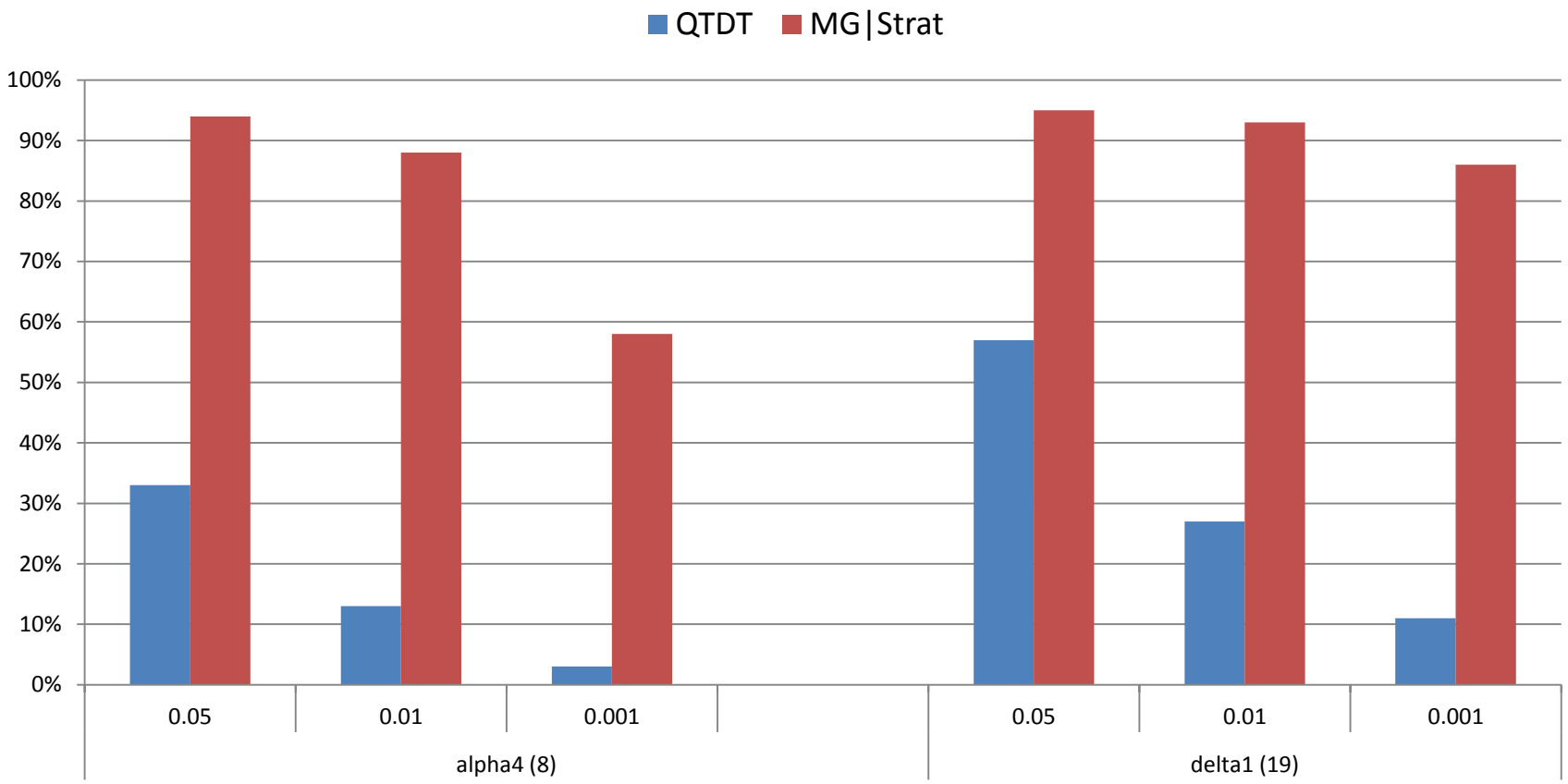
Empirical error rates < nominal values, except for QTDT (2 linked SNPS)
Accounting for pop. Stratification (MG_S) -> decreased mean test statistics
Similar error rates with/out covariate (Diet)
Departure from normality (TG): slight impact on error rates

HDL – Power by P-value



σ^2_{SNP}= 0.3% (DOM) 0.3% (ADD) 0.02% (ADD)

TG – Power by P-value



σ^2_{SNP}= 0.4% (ADD)

0.3% (ADD)

SNP is functional -- Mean χ^2 statistics (μ-χ^2)

			μ-χ^2 (sd)				
Symbol	**SNP**	**Trait**	**QTDT**	**MG**	**MG_S**	**No Ind. with genotype & phenotype data**	**Ne(QTDT)/ N**
alpha4	rs3200218	HDL	17.88 (6.28)	30.96 (8.24)	27.88 (11.55)	5854	32%
alpha2	rs8103444	HDL	1.38 (1.35)	9.56 (4.29)	8.62 (5.05)	5995	37%
delta1	rs8192719	HDL	7.13 (3.8)	17 (5.79)	16.9 (6)	5995	37%
alpha4	rs3200218	TG	2.21 (2.46)	10.93 (5.54)	9.92 (6.31)	5854	32%
		TG_Rob	3.35 (3.16)	13.04 (5.85)	12.67 (6.34)		
delta1	rs8192719	TG	3.11 (2.87)	12.91 (5.28)	12.13 (6.16)		
		TG_Rob	5.15 (3.58)	18.21 (5.89)	17.46 (7.04)	5995	37%

- Power is lowest for
  - Functional SNP with smallest effects (alpha2)
  - Less heritable trait (TG)
  - Non-normal trait (Untransformed vs Transformed TG)

- Mean chi-square QTDT is 1.6 to **6.2** times lower than that of MG_S
  Consistent with the amount of data used (Ne)

Extensions: **Whole-genome association study ?**

- MG adjusted for pop stratification
  - SNP by SNP (candidate gene study)
  - Genomic kinship (Aulchenko et al.,); PCA in family data?

Extensions: **Can association explain linkage ?**

- Major gene, polygenes, environment

$$\Omega_{ij} = \begin{cases} \sigma^2_a + \sigma^2_g + \sigma^2_e & i=j \\ \mathbf{2\varphi_{marker(ij)}\sigma^2_a} + 2\,\phi_{ij}\,\sigma^2_g + \sigma^2_e & i \neq j \end{cases}$$

- Fine-Mapping:
  Are there other associated alleles to be found ?