



RUG



Medical Genetics

University of
Groningen

The Netherlands

Testing for association in whole genome screens

Gerard te Meerman

Dept Genetics

University Medical Center Groningen



RUG



Medical Genetics

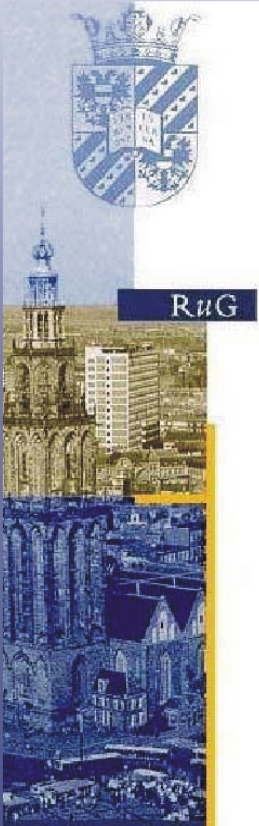
University of
Groningen

The Netherlands

Indirect observation of genes

- Whole genome screens capture the genomic variation only partially
- If marker alleles differ in frequency with nearby causative alleles, correlations may be weak
- Haplotypes capture hidden variation because of the limited number of founders principle

How many founders

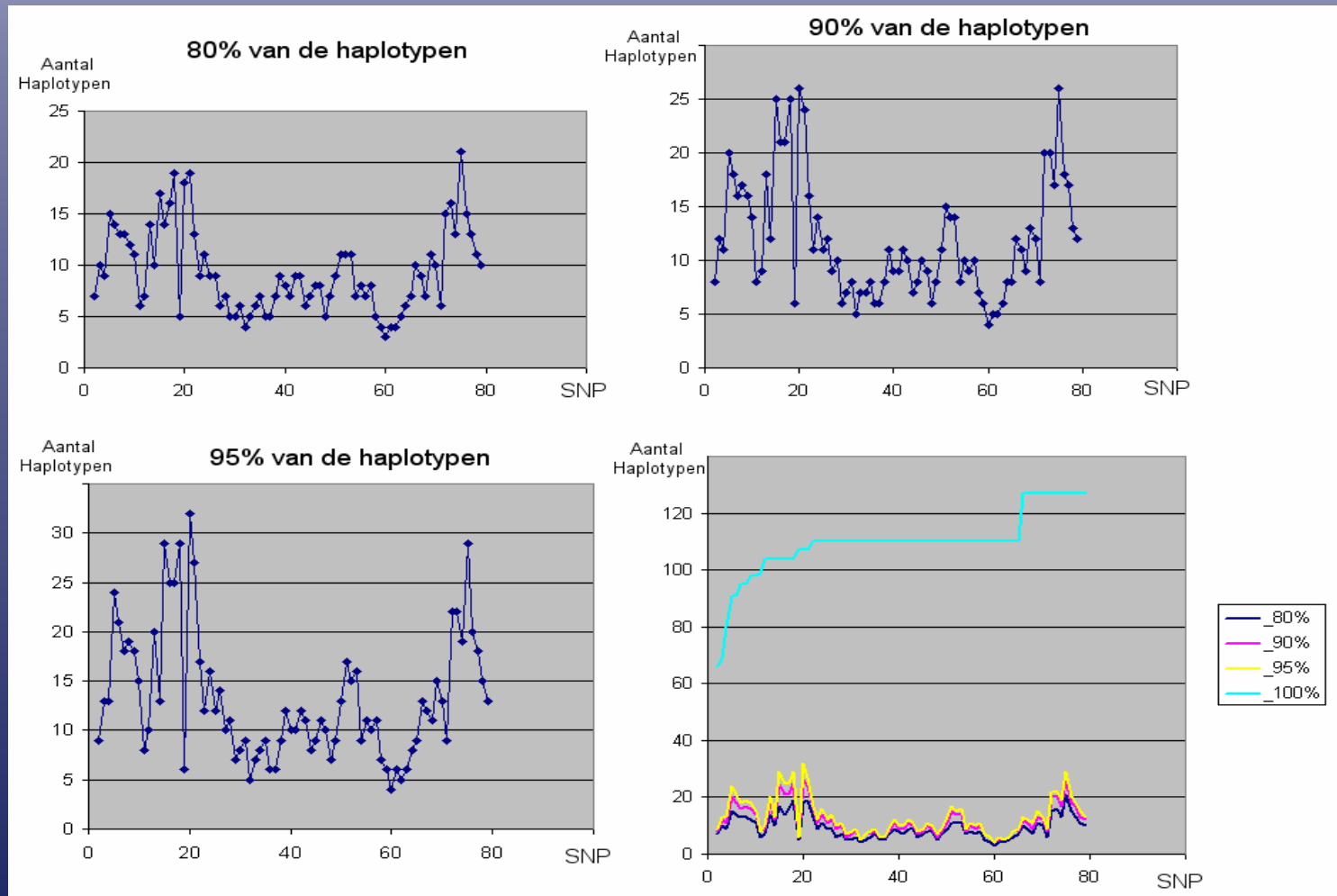


RuG

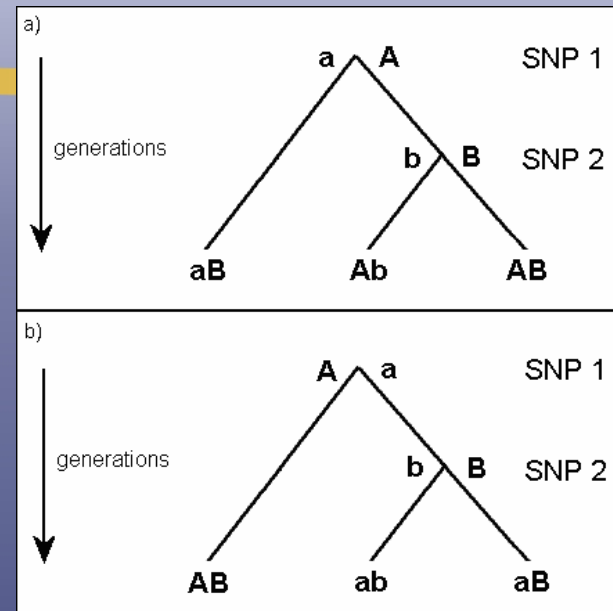
Medical Genetics

University of Groningen

The Netherlands



$3/4$ rule



Hudson and Kaplan. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 111, 147-164 (1985).



RUG

Medical Genetics

University of
Groningen

The Netherlands

Conservation of history



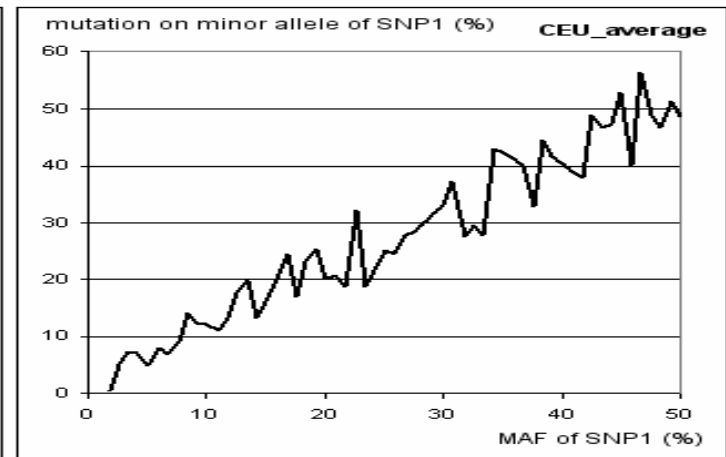
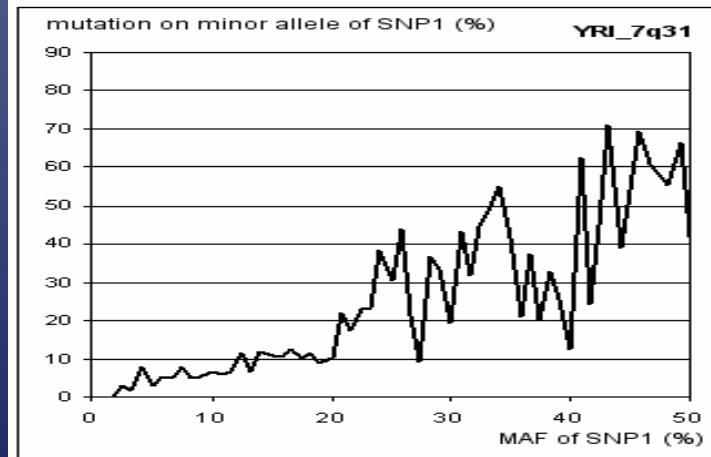
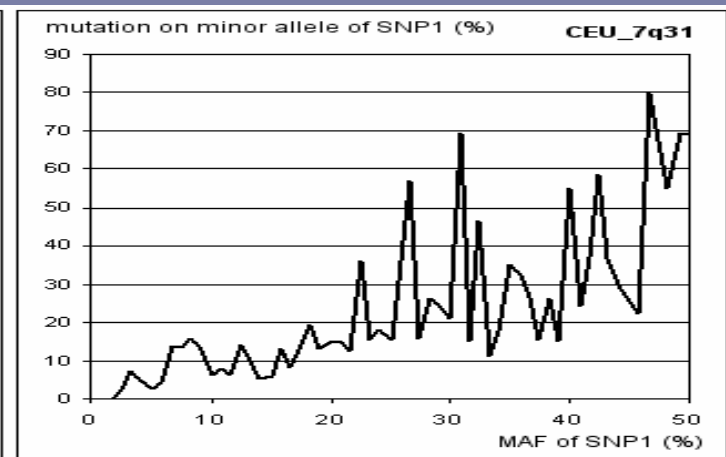
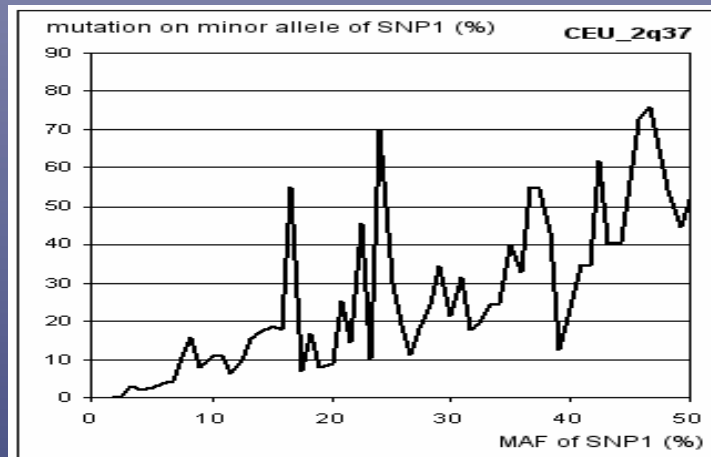
RUG



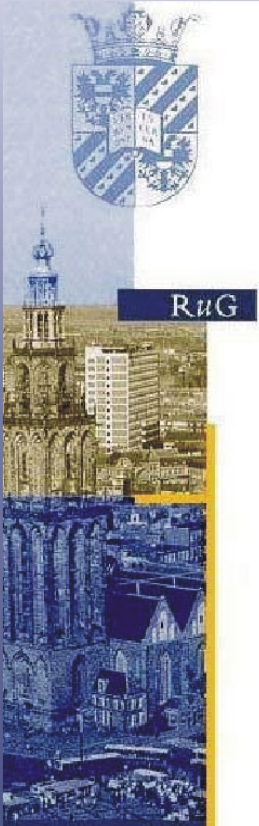
Medical Genetics

University of
Groningen

The Netherlands



$3/4$ rule and recombination rate

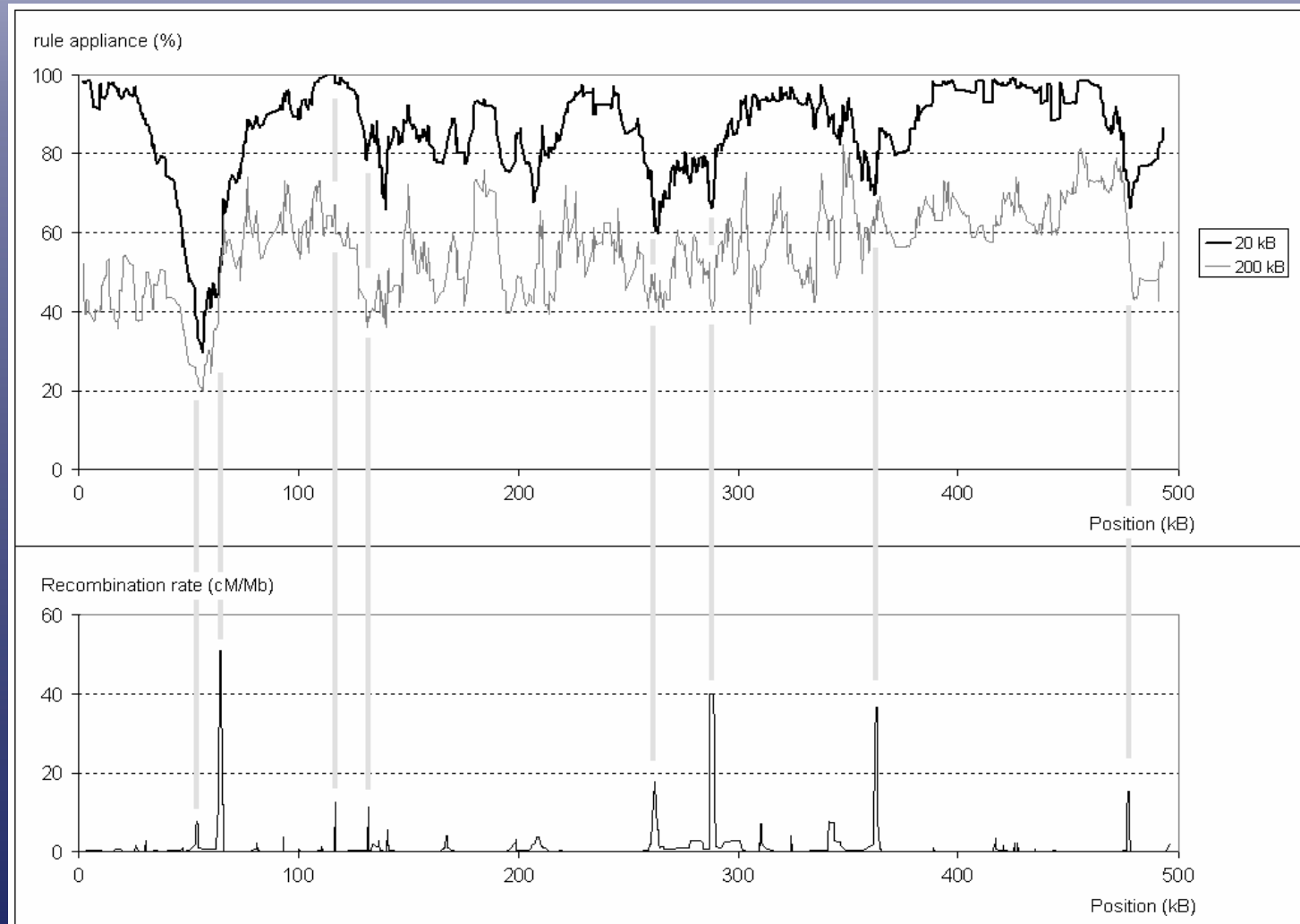


RuG

Medical Genetics

University of
Groningen

The Netherlands





RUG



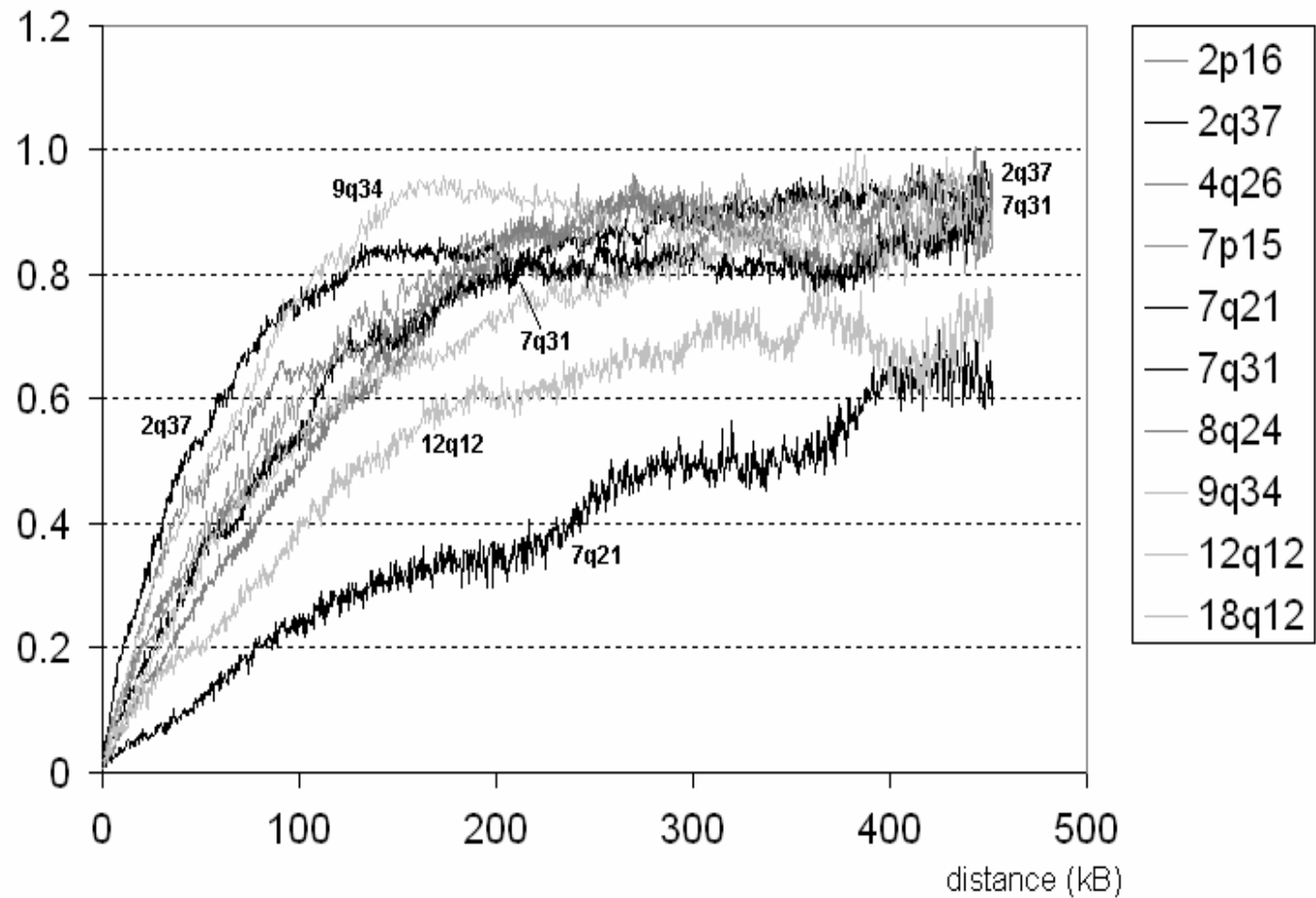
Medical Genetics

University of Groningen

The Netherlands

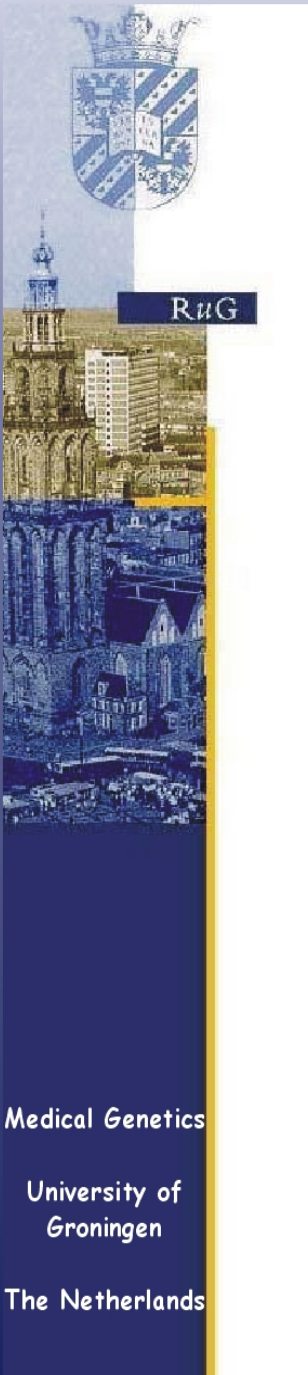
$\frac{3}{4}$ rule and distance

observed FLFC / expected



Information in haplotype length

- We (Martin van der Meulen, Geert Spijker, Ilja Nolte, Andre de Vries) have introduced sharing statistics where length of shared haplotype is used as statistic
- This measures time to coalescence differences that do not need to reflect in association, especially for recent mutations



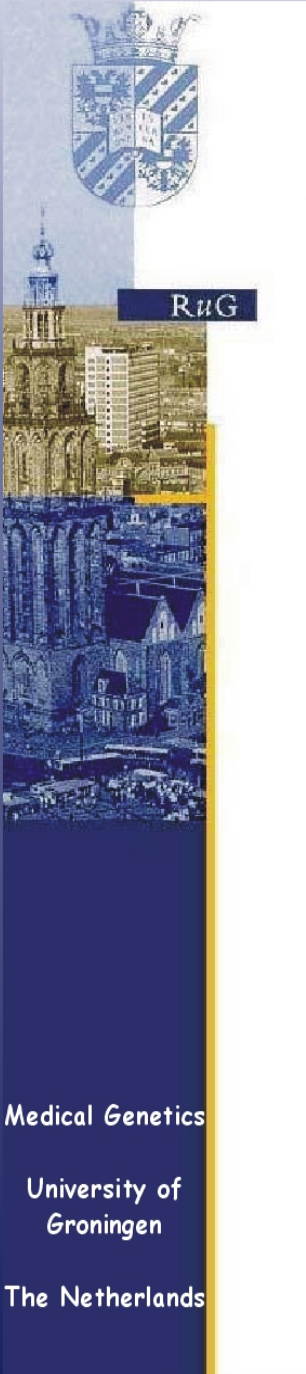
Medical Genetics

University of
Groningen

The Netherlands

Possible statistics

- Compare length within patient haplotypes to length within control haplotypes (HSS test)
- Compare length of patient+control haplotypes to patient x control (Cross test)
- These tests are independent and can be combined
- Evaluation by U statistics and randomization tests.



Medical Genetics

University of
Groningen

The Netherlands

Note

- We need methods to deal rationally with several partially dependent statistical tests
- The dependency structure may not be uniform –leading to ad-hoc methods
- The most disturbing aspect of our work for non-statisticians is the ability to find new, hardly different tests
- Our task is to convince MD's that you need more and better data instead of marginally improved statistical methods



RUG

Medical Genetics

University of
Groningen

The Netherlands

Design issues

- Trio's versus case controls
- Phase determination method of Dimitru Brinza and Alexander Zelikovsky 2SNP: scalable phasing based on 2-SNP haplotypes (and variants)
- Speed of processing : MCMC and E-M are too slow and cannot be used except selectively
- Sequential Randomization and U statistics are fast enough



RUG

Medical Genetics

University of
Groningen

The Netherlands

Pro and con's of trio's

- Better control of quality
- Less control of untransmitted alleles
- Very high quality of genotyping is required (but is no problem anymore)
- Better control of admixture (especially if child is affected)
- More work, less information ???
- Better haplotypes less inflated statistics



RUG

Medical Genetics

University of
Groningen

The Netherlands

How to interpret P values

- There is no a priori hypothesis unlike in Mendelian disease
- We need therefore overwhelming evidence to overcome multiple testing (Bonferroni) and skepticism (Terwilliger, Ioannidis)
- Low P values do however not require extremely large samples (although much larger ones than currently feasible)



RUG

Medical Genetics

University of
Groningen

The Netherlands

Final remarks

- We observe a consistent pattern of haplotypes
- We have reasonably efficient tests
- If nature is not too much misbehaving we can identify gene regions within some 20k
- Tools for finding locations and functional analysis are both indispensable



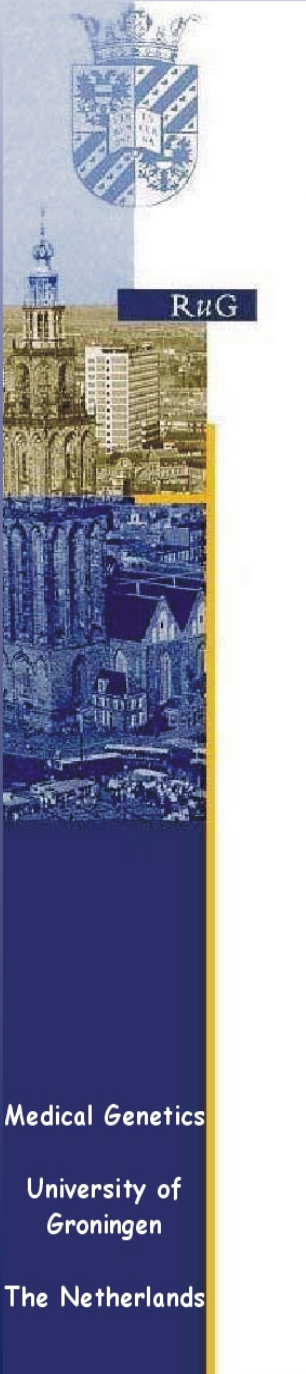
RUG

Medical Genetics

University of
Groningen

The Netherlands

Thanks to my collaborators



• Martin van der Meulen

• Ilja Nolte

• Geert Spijker

• Marijke Niens

• Andre de Vries

• Rudolf Fehrmann

• Paul van der Zwaag

• And the late Lodewijk Sandkuijl