* Yang 2014 first/second local maximum of out of sample prediction accuracy
* Yang 2014: genomic control gactor lamda\_{median}
* Effect size of causal markers: does this just mean they are included in GRM and how do we determine the effect size of the marker in MLM
* Yang 2011: how do we determine the reference allele
  + Seems we only count number of copies of reference allele, but is there not significance wo which non-reference alleles are present not just that they are non-reference alleles
* Candidate vs non candidate markers
  + So in mlm we include which allele of the candidate marker is present as a fixed effect, or also number of copies of reference allele at marker sites as fixed effect?
* Liability scale for heritability h^2
* Yang 2014
  + Selecting M\_T markers by global maximum from out-ofsample prediction accuracy: does that mean looking at all possible subsets and calculating out of sample accuracy for each
* Single best tag: The 24 known associated markers were defined by, for each UC-associated marker listed in the NHGRI GWAS catalogue, a single best tag at *r*2>0.4 from the set of 458,560 if available
* PCA is taking principal components of design matrix? And then including them along with design matrix? Doesn’t that introduce collinearity
* r^2 is the correlation between the x\_j for causal marker and the x\_j for each marker in LD with it?: Equation (10) of ref. 11 states that *Nh r s M mean g mean mean*   1 2 2 / , where *N* is the number of samples, 2*g h* is the heritability explained by *M* genotyped and/or imputed markers, *r*2 for a given causal marker is the average of *r*2 values with all *s* markers in LD with that causal marker, and *r*2 mean and *s*mean denote means across markers (which approximate the means across causal markers)
* yang 2014 supp pg 34: **Effective number of independent markers: what is point of determining this**
* yang 2014 supp pg 35: the mean of the *χ*2 assocation statistics (*λ*mean) from linear regression (LR) analyses: mean over what?
* How does a probe contain a snp (gilad2008 pg 4)
* 
* Seems that GNull is already in use

11-3-18

* From grant proposal: Note that all models are in the “non-estimable” range [23,24]. That is, even asymptotically the numbers of type 1 and type 2 errors do not converge to zero.
  + How do we do this calculation
  + What are the references