Reviewer's report

Title: Combinatorial Mismatch Scan (CMS) for loci associated with dementia in the Amish

Version: 1 Date: 21 November 2005

Reviewer: Mark Abney

Reviewer's report:

General
The difficulties with performing linkage and association analyses in large, complex pedigrees can be very significant. In this manuscript the authors use combinatorial mismatch scanning (CMS) as a way to test for association in individuals affected with dementia from Alzheimer's disease. Although the results are, by themselves, not particularly exciting, the association found with d5s211 may prove interesting in connection with results from other studies. I do have some comments which, I feel, need to be addressed.

-------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

-------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. My understanding of CMS is that it does not compare IBD between cases and controls, rather it looks for differences in either allele or genotype frequency (i.e. IBS) between cases and controls, with power coming from the expectation that IBD sharing in the cases should enrich the allele or genotype associated with the disease. Unless my understanding is faulty, language throughout the manuscript (e.g. the abstract, introduction and methods sections) does not accurately describe this (e.g. page 9 "The first test was for IBD allele frequency inequality...") and should be rewritten more precisely.

2. On page 6, the authors give a coefficient of inbreeding of 0.0151 and state that this is "equivalent to having first cousins as parents." This statement is inconsistent with the quoted inbreeding coefficient in that first cousin matings result in an inbreeding coefficient of 0.0625. Their number is equivalent to having second cousins as parents.

3. In light of the difficulties in interpreting the results of CMS and the p-values obtained, I think it would be useful to have an idea of what the kinship coefficient distributions within and between cases and controls are (not just the means).

4. The authors correctly note that Fisher's exact test may not be entirely valid in that the frequency distributions in cases may not be
the same as in controls due to possibly being more related. What they
do not state is that their empiric method of finding p-values by
permuting affection status suffers from the same problem. In order for
a permutation test to be valid the elements being permuted must be
exchangeable. Under the null that dementia is entirely sporadic with
no genetic component, permuting affection status may be acceptable, but,
under the more realistic null that there is a genetic component, just
not linked to the marker being tested, exchangeability fails. I would
expect that the result would, again, be anticonservative p-values. CMS
attempts to ameliorate this problem by including only more distantly
related individuals, but it is hard to know just how effective this
is. Although I do not believe this invalidates the results, I think
this issue should be addressed and additional relevant information
(i.e. the kinship coefficient distributions mentioned above) would
help provide the proper context to interpret the given p-values.

5. The authors state how many, and which, markers met the p < 0.05 and p
< 0.005 thresholds, but do not mention that the numbers are
essentially exactly what one would obtain solely from false-positives
under the null. I believe the authors should mention this.

Discretionary Revisions (which the author can choose to ignore)

6. If the authors wish to pursue this type of case-control study in the
Amish, I would suggest that a more powerful method than CMS would be
that which is presented in Bourgain et al (2003) Amer J Hum Genet
73:612 or recent, higher powered extensions being developed by MS McPeek.

What next?: Accept after minor essential revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests.