

Parent-of-Origin Effect and Risk for Attention-Deficit/Hyperactivity Disorder: Balancing the Evidence against Bias and Chance Findings

To the Editor: In the December 2005 issue of *The American Journal of Human Genetics*, Hawi et al. made the intriguing observation that a group of catecholamine-related genes, shown elsewhere¹ to be associated with attention-deficit/hyperactivity disorder (ADHD [MIM 143465]), further demonstrate a “consistent pattern of preferential paternal transmission of risk alleles to affected children with ADHD.”^{2(p.959)} The hallmark of this article is that it collated transmission/disequilibrium information from several genes and tested a combined genetic hypothesis, which led to the identification of remarkably significant statistical effects. We believe that the conclusions of this article are flawed, for the following reasons.

1. Hawi et al.² used two alternative criteria to determine whether a gene/locus is associated with ADHD and, therefore, whether it should be retained in the main analysis comparing paternal and maternal transmissions. The first criterion was that association with ADHD must have been demonstrated in their sample of Irish children (significant overtransmission of a specific allele with $P \leq .1$). Of the 17 genes listed in table 1 of their article,² 6 genes (*DRD4* [MIM 126452], *DRD5* [MIM 126453], *TH* [MIM 191290], *DDC* [MIM 107930], *SERT* [MIM 182138], and *TPH2* [MIM 607478]) fulfilled this criterion and were included in the parent-of-origin analysis. Alternatively, Hawi et al. chose genes/loci that “have been confirmed (by several groups) to be associated with ADHD.”^{2(p.960)} With the use of this alternative criterion, three additional genes were retained in the parent-of-origin analysis (*DAT1* [MIM 126455], *SNAP-25* [MIM 600322], and *SHT1B* [MIM 182131]), even though they did not show association with ADHD in the sample presented in their study.² We believe that both criteria are problematic, for the following reasons:
 - a. Had the first criterion been used alone, three genes would have been excluded from the parent-of-origin analysis: *DAT1* ($P = .4$), *SNAP-25* ($P = .12$), and *SHT1B* ($P = .2$). The exclusion of these three genes from the joint analysis of paternal versus maternal transmission would have resulted in a marginal parent-of-origin effect in the remaining six genes ($\chi^2 = 4.07$; $P = .04$). In addition, the claim made by the authors that the lenient threshold ($P \leq .1$) would protect against type II error and would “lead to an underestimate of the size of parent-of-origin effects”^{2(p.960)} is not a valid one, since the statistic used to test association between ADHD and each allele in a given gene/locus (transmission/disequilibrium test [TDT]) and the one used to test for parent-of-origin effect of this allele (χ^2 test with 1 df) are not necessarily correlated. Indeed, it is possible to observe a significant TDT in the absence of parent-of-origin effect (the risk allele is equally overtransmitted from mothers and fathers) and a nonsignificant TDT in the presence of parent-of-origin effect (paternal overtransmission and maternal undertransmission or vice versa, where the effects are canceled out in a global TDT test). Thus, it is not possible to predict the behavior of one statistic given the behavior of the other one, and, consequently, the decision to set the threshold of the individual TDT at $P \leq .1$, to include an allele in the analysis of parent-of-origin effect, is arbitrary. Interestingly—and in contrast to the claim that a lenient threshold of .1 is conservative, with regard to parent-of-origin effect—had the authors chosen a slightly more stringent criterion ($P \leq .07$) without invoking any other criterion, a fourth gene (*SERT*) would have been excluded from the analysis, leading to a nonsignificant statistic of the parent-of-origin effect in the joint analysis of the five remaining genes ($\chi^2 = 1.91$; $P = .17$).
 - b. The second criterion is also problematic, for at least two reasons. First, the literature still lacks consensus on which genes are implicated in ADHD and which are not. An excellent illustration of this problem is provided by the authors themselves. Indeed, in two earlier publications, they reported that *DAT1*³ and *DBH*⁴ contribute significantly to the risk of ADHD. However, both of these associations have not been confirmed in the extended sample presented in the 2005 study.² Second, this criterion seems to reflect a post hoc decision that favors their postulated hypothesis. For example, *DAT1*, which has the most negative effect on sensitivity analysis (P value dropped from .0019 to .013), was “rescued” using this criterion.
2. Remarkably, when we used a χ^2 statistic to compare paternal and maternal transmission of the risk allele separately for each of nine individual genes selected by Hawi et al.,² only two of these alleles in two genes (*DAT1* [$P = .03$] and *SERT* [$P = .009$]) resulted in a significant overtransmission from fathers compared with mothers. Given that the samples used to calculate the χ^2 statistic for each individual allele are quite small (particularly for *SERT* and *DDC*, for which some of the counts are as low as two) and, additionally, that these P values need to be corrected for the large number of tests conducted (at least 17 genes, not to mention the markers in each gene), this is really not an impressive observation and may, in fact, simply reflect chance findings.
3. The likelihood that these results represent chance find-

ings from this data set is clearly illustrated by the simple statistical analysis performed using the data provided in table 1 of Hawi et al.² We correlated the number of markers tested in each gene/locus with the strength of the association between the risk allele in that gene/locus and ADHD, as measured by the χ^2 values. A highly significant correlation was observed (Spearman $R = 0.7$; $P = .002$; $N = 17$).

4. The dramatic contrast between the overtransmission of the parental risk allele at $P = 1.5 \times 10^{-10}$, as opposed to the meager overtransmission of the risk allele from the mother's side ($P = .026$), may also be a reflection of the arbitrary nature of inclusion and exclusion criteria. Indeed, if these criteria were biased in a way in which alleles that are overtransmitted from the paternal side are more likely to enter the parent-of-origin analysis, there would be a highly significant difference between the transmission of parental risk alleles summed over the nine loci compared with the transmission of the same maternal risk alleles summed over the nine loci. This impressive contrast between paternal and maternal overtransmission of the so-called risk allele may be observed even if paternal transmission of the risk allele only marginally exceeds maternal transmission of the risk allele at each individual gene/locus selected by the authors.
5. More generally, results showing significant overtransmission of an allele from one but not the other parent could be interpreted in two different ways. First, the overtransmission from one parent could reflect a true-positive result, and its absence from the other parent could be interpreted by invoking a parent-of-origin effect. Alternatively, the absence of overtransmission from one parent could reflect a true-negative result, and its presence from the other parent could be interpreted as a false-positive result. Although Hawi et al.² systematically sided with the first interpretation, we believe that the second interpretation should be carefully considered before retaining the first one, for several reasons. First, as mentioned above, this particular data set is likely to contain false-positive results. Second, none of the genes studied by Hawi et al.² is known to be associated with any of the molecular mechanisms underlying parent-of-origin effect (e.g., genomic imprinting and trinucleotide-repeat instability), and, to the best of our knowledge, no other studies have reported parent-of-origin effects in any of these genes in relation to ADHD. Third, it has been recognized that true-positive genetic-association results in complex disorders are rather rare.⁵ Finally, the second interpretation is simpler and does not invoke any complex mechanisms such as parent-of-origin effect, which makes it more compatible with the principle of parsimony.

For all these reasons, we call into question the validity of the results of the work of Hawi et al.²

RIDHA JOOBER AND SAROJINI SENGUPTA

Web Resource

The URL for data presented herein are as follows:

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for ADHD, *DRD4*, *DRD5*, *TH*, *DDC*, *SERT*, *TPH2*, *DAT1*, *SNAP-25*, and *5HT1B*)

References

1. Hawi Z, Kirley A, Lowe N, Fitzgerald M, Gill M (2003) Recent genetic advances in ADHD and diagnostic and therapeutic prospects. *Expert Rev Neurother* 3:453–464
2. Hawi Z, Segurado R, Conroy J, Sheehan K, Lowe N, Kirley A, Shields D, Fitzgerald M, Gallagher L, Gill M (2005) Preferential transmission of paternal alleles at risk genes in attention-deficit/hyperactivity disorder. *Am J Hum Genet* 77:958–965
3. Gill M, Daly G, Heron S, Hawi Z, Fitzgerald M (1997) Confirmation of association between attention deficit hyperactivity disorder and a dopamine transporter polymorphism. *Mol Psychiatry* 2:311–313
4. Daly G, Hawi Z, Fitzgerald M, Gill M (1999) Mapping susceptibility loci in attention deficit hyperactivity disorder: preferential transmission of parental alleles at *DAT1*, *DBH* and *DRD5* to affected children. *Mol Psychiatry* 4:192–196
5. Ioannidis JP (2005) Why most published research findings are false. *PLoS Med* 2:e124

From the Departments of Psychiatry (R.J.), Neurology and Neurosurgery (R.J.), and Human Genetics (R.J.; S.S.), McGill University, and Douglas Hospital Research Centre (R.J.; S.S.), Montreal

Address for correspondence and reprints: Dr. Ridha Joobar, Douglas Hospital Research Centre, 6875 Boulevard LaSalle, Montréal H4H 1R3, Québec, Canada. E-mail: ridah.joobar@douglas.mcgill.ca

Am. J. Hum. Genet. 2006;79:765. © 2006 by The American Society of Human Genetics. All rights reserved.

0002-9297/2006/7904-0021\$15.00

Reply to Joobar and Sengupta

To the Editor: We are grateful for the opportunity to respond to the letter by Dr. Ridha Joobar and Dr. Sarojini Sengupta,¹(in this issue) which criticizes our article² published in the December 2005 issue of *The American Journal of Human Genetics*. We believe that the correspondents misunderstood our analysis and have misrepresented our strategy and findings.

First, Joobar and Sengupta contend (point 1b in their letter¹) that there is insufficient evidence of association of any genes with attention-deficit/hyperactivity disorder (ADHD [MIM 143465]) and that our inclusion criteria are therefore arbitrary. Contrary to the authors' opinion, there is a general consensus that *DRD4* (MIM 126452), *DRD5* (MIM 126453), and *DAT1* (MIM 126455) contribute to the development of ADHD, although the odds ratios (ORs) are small and biological mechanisms have not been established. In fact, Dr. Joobar and colleagues reviewed the literature and concluded that the association of *DAT1* and *DRD4* with ADHD "appears to be one of the most repli-

cated in psychiatric genetics and strongly suggests the involvement of the brain dopamine systems in the pathogenesis of ADHD.^{13(p.27)} Meta-analyses of published data for *DRD4* and *DAT1* support this conclusion,^{4,5} and a joint analysis involving 1,980 probands with ADHD and 3,072 of their parents showed association with the *DRD5* locus.⁶

Second, we address their criticism of our inclusion criteria (points 1a and 4 in their letter¹). Alteration of our inclusion criteria to exclude four genes (*DAT1*, *SNAP-25*, [MIM 600322], *5HT1B* [MIM 182131], and *SERT* [MIM 182138]) is an arbitrary decision of the sort that we are accused of making and is a false demonstration of the sensitivity of the results. We refute the suggestion that our selection criteria were selected post hoc. Comparison of the paternally versus the maternally transmitted risk alleles from all 17 genes (table 1) reveals a significant difference in paternal versus maternal transmission ($\chi^2 = 9.47$; $P = .0021$). Our inclusion criteria were designed in an attempt to further define this effect in the same data set that generated the hypothesis. Indeed, parent-of-origin analysis of the eight excluded genes ($\chi^2 = 1.15$; $P = .284$) suggests an effect specific to genes “most associated”² with ADHD. As we made clear in our original article,² initial informal observation of a paternal trend was the motivation for our analysis.

Now we address the statistical questions raised in point 1a of their letter¹: it is true and relevant that the χ^2 test of paternal versus maternal transmissions is expected to be statistically independent of the association test. Concerning our claim that a lenient genewise threshold of $P < .1$ would reduce type II error and would underestimate the size of parent-of-origin effects, we make the following points:

1. Power to detect such effects admittedly depends both on the threshold and on the magnitude and mechanism of a parent-of-origin effect, and the effect on power in this case is unclear. An excessively low or high threshold will decrease power by dilution or sample-size reduction, respectively, and these factors must be balanced.
2. We expect that a low threshold will dilute the magnitude of any ADHD-specific parent-of-origin effect (because of genes unrelated to ADHD being included in the analysis, which is a scenario Joober and Sengupta¹ feel is likely). On the other hand, a high threshold is, on average, unlikely to change the magnitude of a parent-of-origin effect.

The authors claim (point 3 in their letter¹) that correlation between the number of markers tested at each gene

Table 1. Alleles in All 17 Genes Considered for Our Analysis

Gene	Database Identification	Paternal Alleles					Maternal Alleles					All Alleles				
		No. T	No. NT	χ^2	P	OR	No. T	No. NT	χ^2	P	OR	No. T	No. NT	χ^2	P	OR
Excluded:																
<i>DRD1</i>	rs265981 ^a	10	9	.05	1	1.1	11	8	.47	.65	1.4	28	25	.17	.78	1.1
<i>DRD2</i>	rs1800497 ^a	8	7	.06	1	1.1	7	3	1.6	.34	2.3	16	12	.57	.57	1.3
<i>DRD3</i>	rs6280 ^a	6	6	0	1	1.0	9	7	.25	.84	1.3	21	19	.1	.87	1.1
<i>COMT</i>	rs4680 ^a	38	24	3.2	.1	1.6	27	35	1	.77	.8	82	79	.06	.87	1.0
<i>DBH (TaqI)</i>	rs2519152 ^a	27	24	.18	.78	1.1	31	23	1.2	.34	1.3	77	64	1.2	.31	1.2
<i>5HT2A 102C</i>	rs6313 ^a	22	17	1.64	.5	1.3	26	22	.3	.66	1.2	72	63	.6	.5	1.1
<i>NET</i>	rs5568 ^a	17	5	6.5	.017	3.4	8	10	.2	.82	.8	32	21	2.3	.17	1.5
<i>GRIN2A</i>	rs8049651 ^a	<u>21</u>	<u>21</u>	0	1	1.0	<u>29</u>	<u>27</u>	.07	.89	1.1	<u>69</u>	<u>67</u>	.03	.93	1.0
Total ^b	...	149	113	148	135	397	350
Included:																
<i>DRD4 (-616)</i>	rs12720373 ^a	40	23	4.6	.043	1.17	32	19	3.3	.09	1.7	80	49	7.5	.008	1.6
<i>DRD5 (CA)_n</i>	270166 ^c	57	29	9.1	.0034	2.0	54	36	3.6	.07	1.5	114	69	11.1	.0001	1.7
<i>DAT1 (VNTR)</i>	161500 ^c	33	18	4.4	.048	1.8	23	30	.9	.4	.8	76	63	1.2	.31	1.2
<i>TH (TCAT)_n</i>	180306 ^c	21	12	2.5	.16	1.8	28	19	1.7	.24	1.5	55	35	4.4	.04	1.6
<i>DDC (4-bp ins)</i>	M77828 ^d	7	2	2.8	.18	3.5	10	7	.53	.63	1.4	20	9	4.2	.06	2.2
<i>SNAP-25 (MnII)</i>	rs3746544 ^a	33	22	2.2	.18	1.5	28	33	.4	.6	.8	70	52	2.7	.12	1.3
<i>5HT1B (861G)</i>	rs6296 ^a	36	23	2.9	.11	1.6	29	23	.73	.46	1.3	85	68	1.9	.2	1.3
<i>SERT (D17S1294)</i>	D17S1294	15	2	9.9	.002	7.5	9	10	.05	1.0	.9	26	14	3.6	.08	1.9
<i>TPH2 (rs1843809)</i>	rs1843809 ^a	<u>26</u>	<u>7</u>	10.9	.001	3.7	<u>23</u>	<u>12</u>	3.5	.09	1.9	<u>52</u>	<u>22</u>	12.1	.0006	2.4
Total ^e	...	<u>268</u>	<u>138</u>	<u>236</u>	<u>189</u>	<u>578</u>	<u>381</u>
Grand total^f	...	417	251	384	324	975	731

NOTE.—Sum of paternal and maternal counts do not equal all counts because of the exclusion of trios with two informative parents, where parent-of-origin effect cannot be determined. T = transmitted; NT = not transmitted.

^a dbSNP accession number.

^b Maternal versus paternal transmissions in eight excluded genes: $\chi^2 = 1.15$; $P = .284$.

^c GDB accession number.

^d GenBank accession number.

^e Maternal versus paternal transmissions in nine included genes: $\chi^2 = 9.56$; $P = .0019$.

^f Maternal versus paternal transmissions in all 17 genes: $\chi^2 = 9.47$; $P = .0021$.

and the transmission/disequilibrium test (TDT) statistic is evidence that our finding of a parent-of-origin effect is a chance finding. This is untrue, since the TDT and parent-of-origin statistics are not correlated. It is our view that, even if the associations at these genes amount to type I error, this situation should have no bearing on the comparison of maternal and paternal transmissions.

In agreement with Jooper and Sengupta, but for different reasons, we also find it remarkable that two of the genes demonstrate significant parent-of-origin effects when analyzed individually (point 2 in their letter¹), since the effects are more significant than would be expected by chance. These tests are amenable to a Bonferroni correction for the number of genes. The tests for association at each gene also require this correction, as well as correction for the number of markers at each gene. The latter has not been performed and is complicated by our candidate-gene strategy of pursuing initial findings with extra markers across the gene. We presented the statistics for individual genes, to facilitate exploration of the data by readers.

The authors' alternative explanations of the results (point 5 in their letter¹) are equally invalid. Our results indicate that joint transmissions of risk alleles from each parent separately are significant. Our subsequent test was for a differential rate of overtransmission. The probability of a false-positive finding is given by the *P* value—in this case, .0019—which does not require adjustment for multiple testing, is independent of the significance of the TDT statistics for individual genes, and should not be influenced by the likelihood of potential biological explanations.

The authors' comments¹ regarding molecular mechanisms suggest that they did not carefully read our article.² We draw attention to the paragraph containing this sentence: "Since ADHD-associated genes map to many different chromosomes, it is unlikely, a priori, that all these genes are imprinted."^{2(p.963)} Further experiments may help to clarify whether we are observing a true effect, a methodological bias, or a chance finding. The possibility remains that there is a nonmolecular phenomenon, such as selective mating for genetically influenced ADHD-related traits in the male lineage.

Finally, our work is as we described in our article,² and the suggestion that our selection criteria reflect a post hoc decision that favored the hypothesis is untrue and unwarranted. We welcome suggestions for further tests to confirm or to invalidate our findings, including exploration of criteria for inclusion of genes, and we look forward to seeing our hypotheses tested in independent ADHD and control samples.

RICARDO SEGURADO, ZIARH HAWI, AND MICHAEL GILL

Web Resources

The accession numbers and URLs for data presented herein are as follows:

dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/> (see table 1 for accession numbers)

GDB Human Genome Database, <http://www.gdb.org/> (for *DRD5* [CA]_n [accession number 270166], *DAT1* [VNTR] [accession number 161500], and *TH* [TCAT]_n [accession number 180306])

GenBank, <http://www.ncbi.nlm.nih.gov/Genbank/> (for *DDC* [4-bp ins] [accession number M77828])

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for ADHD, *DRD4*, *DRD5*, *DAT1*, *SNAP-25*, *5HT1B*, and *SERT*)

References

1. Jooper R, Sengupta S (2006) Parent-of-origin effect and risk for attention-deficit/hyperactivity disorder: balancing the evidence against bias and chance findings. *Am J Hum Genet* 79: 765–766 (in this issue)
2. Hawi Z, Segurado R, Conroy J, Sheehan K, Lowe N, Kirley A, Shields D, Fitzgerald M, Gallagher L, Gill M (2005) Preferential transmission of paternal alleles at risk genes in attention-deficit/hyperactivity disorder. *Am J Hum Genet* 77:958–965
3. DiMaio S, Grizenko N, Jooper R (2003) Dopamine genes and attention-deficit hyperactivity disorder: a review. *J Psychiatry Neurosci* 28:27–38
4. Faraone SV, Doyle AE, Mick E, Biederman J (2001) Meta-analysis of the association between the 7-repeat allele of the dopamine D(4) receptor gene and attention deficit hyperactivity disorder. *Am J Psychiatry* 158:1052–1057
5. Maher BS, Marazita ML, Ferrell RE, Vanyukov MM (2002) Dopamine system genes and attention deficit hyperactivity disorder, a meta-analysis. *Psychiatr Genet* 12:207–215
6. Lowe N, Kirley A, Hawi Z, Sham P, Wickham H, Kratochvil CJ, Smith SD, et al (2004) Joint analysis of the *DRD5* marker concludes association with attention-deficit/hyperactivity disorder confined to the predominantly inattentive and combined subtypes. *Am J Hum Genet* 74:348–356

From the Department of Psychiatry and Genetics, Trinity College, Dublin (R.S.; Z.H.; M.G.); and Biostatistics and Bioinformatics Unit, Department of Psychological Medicine, Cardiff University, Heath Hospital, Cardiff (R.S.)

Address for correspondence and reprints: Dr. Ziarh Hawi, Department of Psychiatry and Genetics, Trinity College, Dublin 2, Ireland. E-mail: zhhawi@tcd.ie

Am. J. Hum. Genet. 2006;79:765. © 2006 by The American Society of Human Genetics. All rights reserved.
0002-9297/2006/7904-0022\$15.00

Colorectal Cancer Risk in Monoallelic Carriers of *MYH* Variants

To the Editor: Since low-frequency variants in the base-excision-repair gene *MYH* (MIM 604933) were first demonstrated to confer a recessive colorectal cancer (CRC) risk,^{1,2} there has been speculation of an additional dominant effect.^{3,4} In a recent article, Farrington et al.⁵ described the results of screening a series of 2,239 CRC cases and 1,845 controls for germline variants in the human homolog of the *Escherichia coli muty* gene (*MYH*). In whites, Y165C and G382D are the principal disease-caus-

Table 1. Summary of Published Case-Control Studies of the Relationship between *MYH* Variants Y165C and G382D and Risk of CRC

Study and Ethnicity	Place of Study	<i>n</i>	Cases		Controls		
			No. (%) of Monoallelic Carriers	No. (%) of Biallelic Carriers	<i>n</i>	No. (%) of Monoallelic Carriers	No. (%) of Biallelic Carriers
Croitoru et al. ^{3,a}	Ontario	1,238	32 (2.6)	9 (.7)	1,255	21 (1.7)	0 (0)
Farrington et al. ^{5,b}	Scotland	2,217	46 (2.1)	11 (.5)	1,822	28 (1.5)	0 (0)
Enholm et al. ^{8,c}	Finland	1,003	5 (.5)	4 (.4)	424	0 (0)	0 (0)
Kambara et al. ^{9,d}	Brisbane, Australia	92	2 (2.2)	0 (0)	53	1 (1.9)	0 (0)
Wang et al. ^{10,e}	Minnesota	444	10 (2.3)	2 (.5)	313	4 (1.3)	0 (0)
Peterlongo et al. ^{11,f}							
All ethnicities	New York	555	4 (.7)	2 (.4)	918	7 (.8)	0 (0)
White	New York	244	4 (1.6)	2 (.8)	366	6 (1.6)	0 (0)
Jewish	New York	266	0 (0)	0 (0)	450	0 (0)	0 (0)
Zhou et al. ^{12,g}	Sweden	438	6 (1.4)	0 (0)	469	3 (.6)	0 (0)
Present study:							
White	United Kingdom	2,561	53 (2.1)	4 (.2)	2,695	57 (2.1)	0 (0)

^a Cases were aged 20–74 years; study included age- and sex-matched population controls.

^b Of the cases, 872 were <55 years old at diagnosis; study included age- and sex-matched population-based controls.

^c Study included unselected cases (mean ± SD age at diagnosis 67.2 ± 12.1 years) and blood-donor controls.

^d Mean ± SD age of cases was 69.1 ± 10.6 years; study included blood-donor controls.

^e Of the cases, 116 were <50 years old at diagnosis; controls were taken from among individuals undergoing screening colonoscopy with no evidence of adenomatous polyps.

^f Mean age of cases was 62.2 years; study included population-based controls matched by age (mean age 55.0 years), ethnicity, and religious group.

^g Study included sporadic CRC cases and blood-donor controls.

ing variants of *MYH*. Among the cases, Farrington et al.⁵ detected 46 monoallelic carriers of these variants (14 Y165C heterozygotes and 32 G382D heterozygotes) and 11 biallelic carriers (8 G382D homozygotes and 3 compound heterozygotes), with corresponding frequencies among the controls of 28 monoallelic carriers and 0 biallelic carriers. They employed the method of Hugot et al.⁶ to estimate the genotype relative risk (GRR) associated with biallelic and monoallelic variant carriers. Their study confirmed the well-established increased risk of CRC in biallelic carriers of variants in *MYH* (GRR = 92.65; 95% CI 41.60–213.20). Although they did not find a statistically significant increased risk for monoallelic variant carriers (GRR = 1.35; 95% CI 0.92–2.07), the authors suggested that *MYH* heterozygosity might be associated with an elevated CRC risk—primarily in later life—after they arbitrarily restricted their analysis to cases of CRC diagnosed after age 55 years (GRR = 1.68; 95% CI 1.07–2.95).

We believe that such a conclusion may be premature. First, the GRR calculation approach employed is not necessary for assessment of risk associated with heterozygote variant carriers. Moreover, other analytical approaches, including the standard asymptotic approach and exact approaches, yield 95% CIs that do not exclude an odds ratio (OR) of 1. Second, the type of stratification employed raises the issue of post hoc analysis. Third, we have similarly determined the frequencies of Y165C and G382D in a large case-control study and find no evidence that monoallelic *MYH* variant status influences CRC risk.

Our analysis was based on a series of 2,561 patients with histologically confirmed colorectal adenocarcinomas

(1,474 males and 1,087 females; mean ± SD age at diagnosis 61 ± 11.4 years) ascertained through an ongoing initiative at the Institute of Cancer Research/Royal Marsden Hospital National Health Service Trust. We previously reported *MYH* results for a subset of 358 of these cases.⁴ A total of 2,695 control individuals (836 males and 1,859 females; mean ± SD age 59 ± 10.9 years) were the spouses of patients with malignancies, recruited as part of the National Cancer Research Network Trial (1999–2002), the Royal Marsden Hospital Trust/Institute of Cancer Research Family History and DNA Registry (1999–2004), the National Study of Colorectal Cancer Genetics Trial (2004), or the UK Study of Breast Cancer Genetics, all established within the United Kingdom. None of the controls had a personal history of malignancy at the time of ascertainment. All cases and controls were British whites, and there were no obvious differences in the demography of cases and controls in terms of place of residence within the United Kingdom. Blood samples were obtained with informed consent and ethics review board approval, in accordance with the tenets of the Declaration of Helsinki. Genotyping of Y165C and G382D was performed using customized Illumina Sentrix Bead Arrays in accordance with the manufacturer's protocols. Assay validation was conducted using TaqMan and by direct sequencing of a subset of samples.

Among the patients with CRC, we identified 4 individuals with biallelic variants (1 G382D homozygote and 3 compound heterozygotes) and 53 with monoallelic variants (38 G382D heterozygotes and 15 Y165C heterozygotes). Among controls, no biallelic variants were iden-

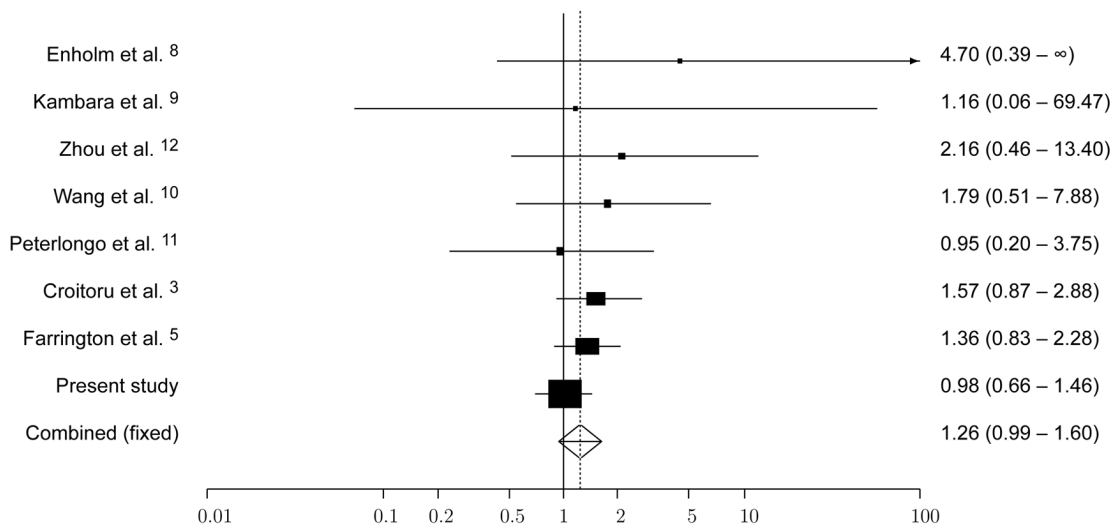


Figure 1. Funnel plot of OR of CRC risk associated with monoallelic Y165C and G382D *MYH* variants, under a fixed-effects model. Studies are plotted in order of decreasing variance of the log(OR). Horizontal lines represent 95% CIs. Each box represents the OR point estimate, and its area is proportional to the weight of the study. The diamond and broken line represent the overall summary estimate, with the 95% CI given by the width of the diamond. The unbroken vertical line is at the null value (OR 1.0).

tified, but 57 monoallelic variant carriers were identified (40 with G382D and 17 with Y165C). For each SNP, genotype distributions among controls did not deviate significantly from Hardy-Weinberg equilibrium (exact $P = 1.0$). These frequencies of *MYH* variants are comparable to those documented in other populations—specifically those from the United Kingdom—based series reported by Farrington et al.⁵ ($P = .17$ for controls; $P = .97$ for cases) (table 1)—and translate to risks of 5.57 (95% CI 0.69–∞) and 0.98 (95% CI 0.66–1.46) associated with biallelic and monoallelic carrier status, respectively. Stratification of the data by 10-year age bands provided no evidence that risk associated with monoallelic carrier status was influenced by age ($P = .13$). Furthermore, after the data were partitioned, as by Farrington et al.,⁵ risks associated with early- and late-onset disease were comparable (for age at onset ≤ 55 years, OR 0.83; 95% CI 0.42–1.51; for age at onset > 55 years, OR 1.03; 95% CI 0.66–1.58).

To further explore the possibility that monoallelic variant status might affect CRC, we applied a kin-cohort approach to compare risks in the 14,668 first-degree relatives of carriers and noncarriers. Data on history of any type of cancer, including age at diagnosis as well as vital status and current age or age at death, were collected for parents, siblings, and offspring by a previously validated questionnaire. Fourteen (4.3%) of the 324 relatives of variant carriers and 431 (3.0%) of the 14,344 relatives of noncarriers had received a diagnosis of CRC. Age-specific cumulative CRC distributions in first-degree relatives were estimated using a marginal-likelihood approach,⁷ and bootstrap estimates for the hazard ratios (HRs) were used to calculate 95% CIs. The HR generated from this analysis for CRC

associated with monoallelic variant status was 1.74 (95% CI 0.62–3.60).

To date, seven published studies have reported the frequency of Y165C and G382D *MYH* variants in CRC cases and controls^{3,5,8–12} (table 1 and fig. 1). Collectively, these provide information on the frequency of *MYH* variants in 8,546 cases and 7,949 controls. To further quantify the risks associated with *MYH* status, we performed a pooled analysis of these published studies with our data. ORs were calculated for each study by use of exact logistic regression, since five of the studies contained < 5 individuals in a single category. Meta-analysis was conducted using standard methods for combining estimates of ORs based on the weighted sum of the log estimates, with the inverse of the variance of the estimate as the weight. An exact conditional-likelihood approach¹³ was used to obtain a 95% CI for the pooled OR. There was no significant evidence of heterogeneity between studies (Cochran's $Q = 3.74$; $P = .81$); however, we used both fixed- and random-effects models to combine study results. Under the fixed-effects model, the pooled OR for monoallelic carrier status was 1.26 (95% CI 0.99–1.60), whereas, under the random-effects model, the pooled OR was 1.24 (95% CI 0.98–1.59). The risk associated with biallelic carrier status is not finite because there is no representation in controls, but an exact approach yields a lower 95% confidence bound of 7.39 for the risk estimate. Alternatively, a naive approach based on the convention of adding 0.5 to each empty cell generates a risk estimate of 6.06 (95% CI 2.02–18.19). Although this analysis provides robust evidence that carriers of biallelic *MYH* variants are at a significantly increased risk of CRC, the data do not indicate a statistically sig-

nificant excess of *MYH* carriers among CRC cases compared with among controls. Our estimate of the risk associated with monoallelic carrier status is, in fact, likely to be inflated, since we restricted our analysis to the pathogenic *MYH* variants Y165C and G382D and since some individuals heterozygous for these variants may carry additional pathogenic variants. Hence, it is likely that additional, apparently heterozygous cases will, in reality, be compound heterozygotes. For example, in the studies by Croitoru et al.³ and Farrington et al.,⁵ some cases heterozygous for Y165C or G382D carried additional rare variants (three cases and one case, respectively). With our analysis adjusted for these observations, the pooled OR associated with monoallelic variant status is 1.23 (95% CI 0.96–1.58).

Our analysis of quantifying the CRC risk associated with carriers of monoallelic *MYH* variants illustrates an inherent problem in studying low-penetrance variants. By definition, such alleles are not associated with large risks. If the population frequency of an at-risk genotype is low (i.e., <2%), then exceptionally large studies are required to estimate precisely the risks. For example, to detect comprehensively the relative risk of 1.2 would require 22,000 cases and 22,000 controls. In conclusion, we believe that the assertion that monoallelic carrier status for *MYH* variants confers an elevated risk of CRC is unsupported on the basis of current data.

EMILY L. WEBB, MATHEW F. RUDD,
AND RICHARD S. HOULSTON

Web Resource

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for *MYH*)

References

1. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR, Cheadle JP (2002) Inherited variants of *MYH* associated with somatic G:C→T:A mutations in colorectal tumors. *Nat Genet* 30:227–232
2. Sieber OM, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips RK, Bisgaard ML, Orntoft TE, Aaltonen LA, Hodgson SV, Thomas HJ, Tomlinson IP (2003) Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in *MYH*. *N Engl J Med* 348:791–799
3. Croitoru ME, Cleary SP, Di Nicola N, Manno M, Selander T, Aronson M, Redston M, Cotterchio M, Knight J, Gryfe R, Gallinger S (2004) Association between biallelic and monoallelic germline *MYH* gene mutations and colorectal cancer risk. *J Natl Cancer Inst* 96:1631–1634
4. Fleischmann C, Peto J, Cheadle J, Shah B, Sampson J, Houlston RS (2004) Comprehensive analysis of the contribution of germline *MYH* variation to early-onset colorectal cancer. *Int J Cancer* 109:554–558
5. Farrington SM, Tenesa A, Barnetson R, Wiltshire A, Prendergast J, Porteous M, Campbell H, Dunlop MG (2005) Germline

susceptibility to colorectal cancer due to base-excision repair gene defects. *Am J Hum Genet* 77:112–119

6. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G (2001) Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 411:599–603
7. Chatterjee N, Wacholder S (2001) A marginal likelihood approach for estimating penetrance from kin-cohort designs. *Biometrics* 57:245–252
8. Enholm S, Hienonen T, Suomalainen A, Lipton L, Tomlinson I, Karja V, Eskelinen M, Mecklin JP, Karhu A, Jarvinen HJ, Aaltonen LA (2003) Proportion and phenotype of *MYH*-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients. *Am J Pathol* 163:827–832
9. Kambara T, Whitehall VL, Spring KJ, Barker MA, Arnold S, Wynter CV, Matsubara N, Tanaka N, Young JP, Leggett BA, Jass JR (2004) Role of inherited defects of *MYH* in the development of sporadic colorectal cancer. *Genes Chromosomes Cancer* 40:1–9
10. Wang L, Baudhuin LM, Boardman LA, Steenblock KJ, Petersen GM, Halling KC, French AJ, Johnson RA, Burgart LJ, Rabe K, Lindor NM, Thibodeau SN (2004) *MYH* mutations in patients with attenuated and classic polyposis and with young-onset colorectal cancer without polyps. *Gastroenterology* 127:9–16
11. Peterlongo P, Mitra N, Chuai S, Kirchoff T, Palmer C, Huang H, Nafa K, Offit K, Ellis NA (2005) Colorectal cancer risk in individuals with biallelic or monoallelic mutations of *MYH*. *Int J Cancer* 114:505–507
12. Zhou XL, Djureinovic T, Werelius B, Lindmark G, Sun XF, Lindblom A (2005) Germline mutations in the *MYH* gene in Swedish familial and sporadic colorectal cancer. *Genet Test* 9:147–151
13. Martin DO, Austin H (2000) An exact method for meta-analysis of case-control and follow-up studies. *Epidemiology* 11:255–260

From the Section of Cancer Genetics, Institute of Cancer Research, Sutton, United Kingdom

Address for correspondence and reprints: Dr. Richard S. Houlston, Section of Cancer Genetics, Brookes Lawley Building, Institute of Cancer Research, Sutton, Surrey, SM2 5NG, United Kingdom. E-mail: Richard.houlston@icr.ac.uk
Am. J. Hum. Genet. 2006;79:765. © 2006 by The American Society of Human Genetics. All rights reserved.
0002-9297/2006/7904-0023\$15.00

Reply to Webb et al.

To the Editor: Biallelic inheritance of *MYH* (MIM 604933) defects has been consistently shown to increase risk of colorectal disease (MIM 608456) in a number of different populations.^{1–3} However, the increase in risk due to monoallelic inheritance is still under debate. Croitoru et al.² alluded to a monoallelic effect, presenting indirect evidence of nonrandom loss of heterozygosity of the wild-type alleles in tumors of heterozygous patients, and they also demonstrated an excess of familial clustering of

disease in these patients. Our previously published data³ suggested a monoallelic effect, which was statistically significant only for later-onset disease. Rather than a “data-dredging” exercise, the rationale for the analysis of age subgroups was our a priori hypothesis of an age effect. We conducted significance testing by permutation tests, because empirical significance levels are generally considered to be more robust to violations of the underlying statistical assumptions than are the asymptotic significance levels. However, we recognized that this effect was of borderline statistical significance at the 5% level, and we emphasized that this evidence should be interpreted with caution. We concluded that this preliminary observation merited further study.

Since publication of that work in the *Journal*,³ we have performed a replication study, using Scottish population-based samples and a meta-analysis of all published case-control *MUTYH* association studies, and this work was recently published.⁴ The pooled results confirmed the reported biallelic effect and gave more-precise estimates of the associated risk. However, we again observed a monoallelic effect of borderline statistical significance (OR 1.27; 95% CI 1.01–1.61). These findings are comparable to those of the meta-analysis presented by Webb et al.,⁵ with additional primary data from English samples, although their analysis fails to achieve statistical significance at the 5% level (OR 1.26; 95% CI 0.99–1.60). The overall lack of association with the *MUTYH* gene in the Webb study is in contrast to the other two large association studies with >1,000 cases and controls.^{2,3} Similarly, the results of the kin-cohort study performed by Webb et al.⁵ is in contrast to the published work of Jenkins et al.,⁶ who demonstrated a threefold increase in risk for monoallelic carriers by use of a similar analysis. These differences may be due to study bias and confounding due to imperfect case-control matching, rather than to true population differences.

Overall, we think that the available data support a small monoallelic effect of *MYH* variants. However, it is clear that meta-analysis is needed to achieve the very large sample sizes required to confirm the small effects that are typical of such variants. A road map for this effort was recently proposed.⁷ To this end, we have already invited all other seven groups with published case-control data on *MYH* variants to pool all available data, to address this issue and to investigate evidence of age, sex, or other effects associated with tumor pathology.

SUSAN M. FARRINGTON, ALBERT TENESA,
REBECCA BARNETSON, ALICE WILTSHIRE,
JAMES PRENDERGAST, MARY PORTEOUS,
HARRY CAMPBELL, AND MALCOLM G. DUNLOP

Web Resource

The URL for data presented herein is as follows:

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for *MYH* and colorectal disease)

References

1. Sieber OM, Lipton L, Crabtree M, Heinemann K, Fidalgo P, Phillips RK, Bisgaard ML, Orntoft TF, Aaltonen LA, Hodgson SV, Thomas HJ, Tomlinson IP (2003) Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in *MYH*. *N Engl J Med* 348:791–799
2. Croitoru ME, Cleary SP, Di Nicola N, Manno M, Selander T, Aronson M, Redston M, Cotterchio M, Knight J, Gryfe R, Gallinger S (2004) Association between biallelic and monoallelic germline *MYH* gene mutations and colorectal cancer risk. *J Natl Cancer Inst* 96:1631–1634
3. Farrington SM, Tenesa A, Barnetson R, Wiltshire A, Prendergast J, Porteous M, Campbell H, Dunlop MG (2005) Germline susceptibility to colorectal cancer due to base-excision repair gene defects. *Am J Hum Genet* 77:112–119
4. Tenesa A, Campbell H, Barnetson R, Porteous M, Dunlop M, Farrington SM (2006) Association of *MUTYH* and colorectal cancer. *Br J Cancer* 95:239–242
5. Webb EL, Rudd MF, Houlston RS (2006) Colorectal cancer risk in monoallelic carriers of *MYH* variants. *Am J Hum Genet* 79:768–771 (in this issue)
6. Jenkins MA, Croitoru ME, Monga N, Cleary SP, Cotterchio M, Hopper JL, Gallinger S (2006) Risk of colorectal cancer in monoallelic and biallelic carriers of *MYH* mutations: a population-based case-family study. *Cancer Epidemiol Biomarkers Prev* 15:312–314
7. Ioannidis JP, Gwinn M, Little J, Higgins JP, Bernstein JL, Boffetta P, Bondy M, et al (2006) A road map for efficient and reliable human genome epidemiology. *Nat Genet* 38:3–5

From the Colon Cancer Genetics Group, School of Clinical and Molecular Medicine (S.M.F.; A.T.; R.B.; A.W.; J.P.; M.P.; H.C.; M.G.D.), Clinical Genetics Department (M.P.), and Public Health Sciences (H.C.), University of Edinburgh, and Medical Research Council, Human Genetics Unit (S.M.F.; A.T.; R.B.; A.W.; J.P.; M.G.D.), Edinburgh

Address for correspondence and reprints: Dr. Susan M. Farrington, Colon Cancer Genetics Group, University of Edinburgh, 4th Floor, MRC Human Genetics Unit, Crewe Road, Edinburgh EH4 2XU, United Kingdom. E-mail: Susan.Farrington@hgu.mrc.ac.uk

Am. J. Hum. Genet. 2006;79:765. © 2006 by The American Society of Human Genetics. All rights reserved.
0002-9297/2006/7904-0024\$15.00