PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2005) **272**, 643–649
doi:10.1098/rspb.2004.2890
Published online 22 March 2005

Cadherins in maternal–foetal interactions: red queen with a green beard?

Kyle Summers[1]* and Bernard Crespi[2]

[1]*Department of Biology, East Carolina University, Greenville, NC 27858-4353, USA*
[2]*Behavioural Ecology Research Group, Department of Biology, Simon Fraser University, Burnaby, BC, Canada V5A 1 S6*

Cadherins are homophilic cell surface adhesion proteins, some of which mediate interactions between maternal and foetal tissues during mammalian pregnancy. David Haig suggested that these proteins may exhibit 'green-beard gene' effects, whereby the nature of binding between identical alleles in mother and foetus leads to differential levels of resource transfer. The selfish effects of such self-recognizing alleles should, however, be suppressed over evolutionary time by unlinked genes, which is expected to lead to antagonistic coevolution between placentally expressed cadherins and unlinked modifiers. Such molecular coevolution should leave a signature of positive selection, with high ratios of non-synonymous to synonymous amino acid substitution. We present evidence that three placentally expressed cadherin genes, E-cadherin, P-cadherin and VE-cadherin, have been subject to positive selection. By contrast, a 'control' cadherin that is not expressed in the placenta, H-cadherin, showed no evidence of selection. These results provide support for the hypothesis that the cadherin genes involved in maternal–foetal interactions have been subject to green-beard-effect mutations over the course of evolutionary history, leading to antagonistic coevolution with suppressing elements from the parliament of genes.

Keywords: cadherin; selection; placenta; green-beard

1. INTRODUCTION

Maternal–foetal interactions during pregnancy in mammals are extremely intimate, providing the venue for the transfer of nutrients from mother to offspring, and the exchange of cell surface and hormonal signals (Haig 1993). These interactions are expected to involve complex mixtures of cooperation and conflict between the effects of genes expressed in the mother, and the effects of paternal and maternal genes expressed by the foetal genotype (Haig 1993, 1997, 2000; Hurst *et al.* 1996), with important consequences for the biochemical interplay between mother and foetus (Haig & Graham 1991), the risks of pregnancy (Haig 1993, 1999) and the macroevolution of viviparity and placentation among mammals (Crespi & Semeniuk 2004).

Maternal–foetal interactions develop through the invasion and modification of maternal tissues by extra-embryonic foetal tissue (Haig 1993). Many of these interactions involve molecules expressed by the foetal genotype that recognize their counterparts in maternal tissues and *vice versa* (Campbell *et al.* 1995). Haig (1996) proposed that genes coding for cell adhesion molecules involved in maternal–foetal interactions should be prone to mutations leading to 'green-beard' effects. The term 'green-beard gene' was coined by Richard Dawkins (1976) to describe a phenomenon originally conceived by William D. Hamilton (1964). Hamilton envisioned an allele at a single locus (or a 'supergene' involving close linkage of several genes) that caused its bearer to produce a recognizable trait, to recognize that same trait in other individuals and to treat those individuals preferentially. Green-beard effects may play important roles in some types of signalling (e.g. Queller 1984; Guilford 1988). However, these genes have generally been viewed as unlikely to play important roles in recognition systems (Alexander & Borgia 1978) owing to the perceived implausibility that a single gene or supergene could exhibit three such functionally diverse effects.

Recent research indicates that green-beard alleles are more than just a theoretical possibility. In fire ants, specific alleles in a closely linked group of genes apparently allow worker recognition of queen genotype, and mediate genotype-specific aggression (Keller & Ross 1998). More recently, a single locus, the csA gene in the slime mould *Dictyostelium discoideum*, has been demonstrated to mediate green-beard effects. Specifically, the protein (gp80) produced by this gene controls cell–cell adhesion in the context of aggregation, group movement and sporulation in response to starvation (Crespi & Springer 2003; Queller *et al.* 2003). Cells with modified versions of this protein are discriminated against by virtue of the homophilic attraction mediated by the wild-type protein. This example provides confirmation of Haig's (1996) hypothesis of green beards mediated by homophilic cell adhesion molecules. However, the question of whether such molecules exhibit green-beard effects in the context of maternal–foetal interactions remains unexplored.

A key family of molecules involved in cell adhesion during maternal–foetal interactions, and cited as a promising candidate for green-beard effects by Haig (1996), are the cadherins. Cadherins are a class of cell adhesion proteins originally discovered in vertebrates (Hyafil *et al.* 1980). These proteins are ancient components of the animal proteome, and probably evolved before the origin of animals (King *et al.* 2003). A wide variety of cadherins have been identified and characterized, and the phylogenetic relationships among different types have been analysed (Gallin 1998; Nollet *et al.* 2000).

* Author for correspondence (summersk@mail.ecu.edu).

Received 3 June 2004
Accepted 2 August 2004

Cadherins are involved in many fundamental aspects of morphogenesis, cell–cell signalling, and the maintenance of tissue integrity (Edelman *et al.* 1987; Friedlander *et al.* 1989; Klymkowsky & Parr 1995; Wheelock & Johnson 2003). Cadherins are also important in a wide range of cell–cell interactions during implantation, placentation and other maternal–foetal interactions in vertebrates (MacCalman *et al.* 1996; Paria *et al.* 1999; Getsios *et al.* 2000). In particular, some cadherins mediate the development of structures that are necessary for the transfer of nutrients between the mother and the developing embryo in mammals (Zhou *et al.* 1997; Floridon *et al.* 2000; Shih *et al.* 2002).

Haig (1996) proposed that cadherins are likely to be particularly susceptible to the invasion of 'green-beard' mutant alleles, because of their roles in inter-cellular recognition and the transfer of resources from mother to foetus. Thus, cadherins combine two features that make them likely to fulfil the requirements of a green-beard gene: (i) they have modular extracellular domains that recognize and specifically bind to copies of themselves on other cells (Blaschuk *et al.* 1990) and (ii) they have cytoplasmic domains that allow them to influence cellular behaviour. Variant cadherins that expressed preferential self-recognition and interaction across the foetal–maternal interface, resulting in a beneficial outcome for the foetus, would constitute green-beard alleles (Haig 1996).

One important feature of green-beard genes is that they are likely to be in conflict with the interests of other (unlinked) loci in the genome (Alexander & Borgia 1978). If the benefits that accrue to offspring with the green-beard allele are less than the costs imposed on offspring without the green-beard allele, then the net effect of the green-beard allele on unlinked genes is negative (Ridley & Grafen 1981). This kind of conflict is analogous to that seen in systems of meiotic drive, which led Haig (1996) to refer to the evolution of green-beard genes in maternal–foetal interactions as gestational drive. The conflict could be temporary (i.e. if the green-beard allele goes to fixation), but it would nevertheless impose some cost on the 'parliament of genes' (Leigh 1971). In response to this cost, selection would favour modifiers that interfered with the expression of drive (Haig 1996). Such conflicts can lead to antagonistic coevolution, whereby changes in the sequence coding for the green-beard effect that cause the protein to engage in self-preferential interactions are suppressed (later in evolutionary time) by changes in the other, unlinked genes in the genome (Hurst *et al.* 1996). Over long periods of evolutionary time, such intragenomic conflicts may lead to high rates of amino acid change at specific positions that are involved in the interaction of the green-beard protein and its suppressors (i.e. molecular 'red queen' effects). These changes would generate a signature of high non-synonymous substitution rates, relative to rates of synonymous substitution, at key sites in the genes coding for both molecules involved in the interaction (Yang & Bielawski 2000).

In this paper, we tested the hypothesis that cadherin genes expressed in maternal–foetal interactions have been subject to an evolutionary history of diversifying selection for high rates of amino acid change. To do so, we collected DNA sequence data from GenBank on cadherins that are involved in placentation, and we used these sequences to investigate patterns of nucleotide substitution. We also collected sequence data on a cadherin that is apparently not involved in maternal–foetal interactions, as the best form of 'negative control' that is currently available. We predicted that, if green-beard effects have been subject to selection in these genes, then we should see evidence for positive selection, which is a molecular signature of antagonistic coevolution (and other forms of diversifying selection; Yang 2001). We note here that the green-beard hypothesis is not the only hypothesis that could explain the presence of diversifying selection acting on these genes. We discuss alternative hypotheses in the discussion section.

We focused on cadherins known to play crucial roles in maternal–foetal interactions in one or more species of mammals: E-cadherin, P-cadherin and VE-cadherin. Cadherin 6 and cadherin 11 are also known to be expressed in this context, but the numbers of available sequences of these genes were insufficient for analysis.

Several cadherins were identified and characterized in the 1980s (e.g. Gallin *et al.* 1983; Schuh *et al.* 1986). E-cadherin was identified by Takeichi (1991). Like other cadherins, E-cadherin is a transmembrane glycoprotein. It has an extracellular domain consisting of five distinct modules. Each module is approximately 110 amino acids long, and contains specific calcium binding motifs. Four of these modules are highly similar in sequence, whereas a fifth, membrane-proximal module is divergent from the other four (Blaschuk *et al.* 1990). The structure of E-cadherin has been investigated with X-ray crystallographic analysis and nuclear magnetic resonance spectroscopy (Overduin *et al.* 1995; Shapiro *et al.* 1995*a*; Nagar *et al.* 1996; Pertz *et al.* 1999), revealing a seven-stranded beta-barrel structure in each cadherin module. The ectodomain modules form parallel protein dimers at the cell surface, which in turn allow the formation of a zipper-like structure between cadherins on two different cells, maintaining cell–cell adhesion (Tomschy *et al.* 1996). The three-dimensional structure characteristic of E-cadherin is believed to be characteristic of other, closely related cadherins as well (Overduin *et al.* 1995; Shapiro *et al.* 1995*b*). E-cadherin is involved in placentation and invasion of maternal tissues by the trophoblast (Floridon *et al.* 2000), although E-cadherin expression appears to be downregulated during the period in which the trophoblast is most invasive (Shih *et al.* 2002). Recent research indicates that E-cadherin is intimately involved in interaction between the embryonic trophoblast cells and the maternal endometrium (Paria *et al.* 1999). E-cadherin is also crucial for terminal differentiation of the trophoblast (Getsios *et al.* 2000), which could provide an opportunity for green-beard effects, as hypothesized by Haig (1996).

P-cadherin is mainly expressed in the placenta (Nose & Takeichi 1986). Haig (1996) also proposed this cadherin as a candidate for green-beard effects. VE-cadherin is an important modulator of the interaction between the embryonic trophoblast and uterine tissue (Zhou *et al.* 1997). In particular, VE-cadherin is closely involved in cytotrophoblast invasion of the maternal vasculature, and in establishing arterial connections between the embryonic and maternal circulations. This process involves biochemical mimicry by the trophoblast tissues, a process mediated in part via the expression of VE-cadherin (Zhou *et al.* 1997).

This could also provide a context conducive to green-beard effects.

H-cadherin is a member of the neuronal subgroup of cadherins (Nollet *et al.* 2000). It is involved in the development of the heart, and it is not known to be involved in maternal–foetal interactions.

2. METHODS

We searched for DNA sequences for the cadherins involved in maternal–foetal interactions during embryonic development, focusing on multiple sequences from different species, because the methods we used are designed for comparisons among species (Yang 1997). Reference sequences for these regions were obtained via Locuslink on the NCBI website. One reference (from a species in which the particular cadherin is known to be involved in maternal–foetal interactions) was used in a translated protein BLAST (Altschul *et al.* 1997) search (tblastn), in which the protein sequence of the gene is compared with translated sequences from the GenBank database. Homologous sequences with high similarity scores from multiple species were chosen to create a cross-species sequence alignment. For E-cadherin and VE-cadherin, we were able to find sequences from at least four different species of mammals for the analyses (four and six, respectively). For P-cadherin, we could only find sequences from three different species, but we were able to find variants at this locus from mice and humans, bringing the total number of sequences analysed to five.

We constructed multiple alignments from sequences for each gene region: table 1 shows the species used for each gene region, with the GenBank accession number for that sequence. Alignments were created as follows: first, the sequences were copied from the GenBank database and saved in a textfile in FASTA format. This file was analysed with the program RevTrans (Wernersson & Pederson 2003). This program aligns DNA sequences by translating them, aligning the protein sequences and then reverse translating them (maintaining the codon alignment). This process maximizes the probability that the codons in each sequence are properly aligned (i.e. aligned with homologous codons). All alignments used for the analyses are available from the authors on request. The aligned sequences were imported into PAUP 4.0b10 (Swofford 2002). A branch and bound search was used to identify the most parsimonious tree. This tree was then used to calculate the following parameters via maximum likelihood: base frequencies, transition/transversion ratio, gamma parameter (alpha) and proportion of invariant sites. These parameter estimates were then used to parametrize a maximum likelihood analysis of the phylogenetic relationships of the sequences involved. A heuristic search was used, with TBR and 100 random addition-order replicates. The ML tree obtained from PAUP was imported into TreeView (Page 1996) and edited to ensure that it was an unrooted tree appropriate for analysis with PAML (Yang 1997). The tree topologies we found using these methods were consistent with the phylogenetic relationships of the species involved, based on recent analyses of mammalian phylogenetic relationships (reviewed in Springer *et al.* 2004).

Analyses of the ratios of non-synonymous substitution rates to synonymous substitution rates, or d*N*/d*S* ratios (ω) were carried out with PAML (Yang 1997). The sequence alignments from RevTrans were used for the main data file for analyses of positive selection. All analyses used the option Cleandata = 1, which removes gaps in the sequences from analysis. Hence, regions with sequence data from only some species were not analysed. The phylogenetic tree obtained from the PAUP analysis was used as the tree file for a preliminary analysis in PAML using the one-ratio model (M0). The estimates of branch lengths produced by this model were then incorporated into the tree file, and this tree plus branch lengths file was used as the main tree file for all subsequent analyses, as recommended by Yang (1997). Simulation studies indicate that the power and accuracy of the log likelihood ratio tests (LRTs) used to investigate the significance of evidence for positive selection increases with the number and diversity of sequences used (Anisomova *et al.* 2001, 2002). These simulation studies also indicate that the tests are robust if the tree length in the analyses is greater than one substitution per codon. This was the case for the trees used in this study. The following models were used to analyse the dataset for each gene region: M0 (single rate model), M1 (neutral model), M2 (basic selection model), M3 (discrete selection model), M7 (continuous distribution model) and M8 (continuous distribution plus selection model). Codon frequencies were estimated from the average nucleotide frequencies at the three codon positions for all runs, using the F3X4 model (Yang 1997). We did not use branch-specific models in our analyses, given the small size of our datasets. Log LRTs were used to test for significant differences in the fit of the models incorporating selection relative to their counterparts that did not allow positive selection (Yang *et al.* 2000). These tests are employed by calculating two times the difference in log-likelihood between two nested models, and comparing that statistic to a chi-square distribution. The number of degrees of freedom is determined by the difference in the number of parameters estimated in the two models under comparison. We focused on comparing model 1 results to model 2 results, and model 7

Table 1. Gene regions, species and GenBank accession numbers.

gene region	species	GenBank acc. no.
E-cadherin (CDH1)	*Homo sapiens*	NM_004360.2
E-cadherin (CDH1)	*Bos taurus*	AY508164.1
E-cadherin (CDH1)	*Mus musculus*	X06115.1
E-cadherin (CDH1)	*Rattus norvegicus*	NM_031334.1
P-cadherin (CDH3)	*Mus musculus*	AK045041.1
P-cadherin (CDH3)	*Mus musculus*	AK031265.1
P-cadherin (CDH3)	*Homo sapiens*	NM_001793.2
P-cadherin (CDH3)	*Homo sapiens*	X63629.1
P-cadherin (CDH3)	*Rattus norvegicus*	XM_226426.2
VE-cadherin (CDH5)	*Mus musculus*	NM_009868.3
VE-cadherin (CDH5)	*Sus scrofa*	AB046120.1
VE-cadherin (CDH5)	*Bos taurus*	AY363224.1
VE-cadherin (CDH5)	*Homo sapiens*	NM_001795.2
VE-cadherin (CDH5)	*Rattus norvegicus*	XM_226213.2
VE-cadherin (CDH5)	*Pan troglodytes*	AY413035
H-cadherin (CDH13)	*Homo sapiens*	NM_001257.2
H-cadherin (CDH13)	*Mus musculus*	NM_019707.1
H-cadherin (CDH13)	*Rattus norvegicus*	NM_138889.1
H-cadherin (CDH13)	*Pan troglodytes*	AY417901
T-cadherin (CDH13) (Avian H-cadherin)	*Gallus gallus*	M81779.1

to model 8, as recommended by Yang *et al.* (2000). We also compared model 0 to model 3, as this gives an indication of the significance of variation in ω among sites.

The models used differ as follows (Nielsen & Yang 1998; Yang *et al.* 2000): model 0 assumes a single basic rate for all sites. Model 1 includes a category of sites for which $dN = dS$ (neutral evolution), in addition to the category of sites with the basic rate. Model 2 is a discrete model that includes a category of sites under positive selection, such that $dN > dS$, a neutral rate category and a basic rate category. Model 3 uses a discrete distribution to model heterogeneous variation in the ratio of dN to dS, and includes a category of sites for which $dN > dS$. Model 7 uses a beta distribution to approximate a continuous distribution of rates across sites, but does not allow for positively selected sites. Model 8 includes a category of sites with an approximate continuous distribution of rates across sites (as in model 7), but also includes a category of sites for which $dN > dS$. In order to check that the method had not converged on a local maximum (leaving a global maximum undetected) we carried out several runs for each set of models for each gene, using four values (a fixed value of 1 and initial values of 1, 2 and 5) for kappa (transition/transversion ratio), and two values (0.4 and 3.14) for omega (dN/dS ratio). The final likelihoods were compared and the highest likelihoods taken as the best estimate for each case. We did not detect any cases where the initial analyses were trapped on local maxima.

We also used the empirical Bayes method implemented in PAML to estimate the posterior probabilities that specific sites are under positive selection (Yang 1997). The Bayes method allows *post hoc* identification of the specific sites (codons) that are under positive selection.

3. RESULTS AND DISCUSSION

Analyses of E-cadherin, P-cadherin and VE-cadherin each yielded evidence for diversifying selection at the molecular level (table 2 and Electronic Appendix). Both the discrete approximation selection model (M3) and the continuous distribution approximation selection models (beta and ω: M8) showed high dN/dS ratios (ω) in the selection category for E-cadherin. The simplest selection model (M2) did not show a category under positive selection. However, this model is constrained by necessity of including a category of neutral sites, and hence may not include a category of sites under positive selection, even when one exists. Hence, model 3 and model 8 are more reliable in discovering the presence of sites under positive selection (Ziheng Yang, personal communication). In this case, although both models indicate positive selection, the distribution of selection differs under the two models (table 2 and Electronic Appendix). Model 3 indicates a high proportion of sites (approximately 17%) under mild positive selection ($\omega = 1.2$). In contrast, model 8 showed a small proportion of sites ($<1\%$) under extremely strong positive selection ($\omega = 89.4$). This extreme estimate of positive selection is unlikely to be highly accurate, but is likely to reflect the presence of very strong positive selection at one or a few sites (Ziheng Yang, personal communication). The empirical Bayes method reflected the difference between the two models, with significant probabilities of positive selection for 40 different sites under model 3, and only a few sites with high probabilities of positive selection under model 8 (Electronic Appendix). The total number of sites with high probabilities of positive selection was 112 for model 3, but only 6 under model 8. This difference probably reflects the ability of the beta distribution approximation used in the continuous approximation (model 8) to account for variation in the dN/dS ratio that is not well-accommodated by the discrete approximation model. Nevertheless, it is clear that at least two sites in this protein (sites 178 and 703) are under strong positive selection. Log LRTs were highly significant in both comparisons recommended by Yang *et al.* (2000): M2 versus M1 and M8 versus M7. Given that M2 did not identify any sites under selection, the relevant comparison is M8 versus M7, which gives a *P*-value of less than 0.01 with two degrees of freedom when compared with a chi-square distribution. The comparison of M3 with M0 was also highly significant. All things considered, there is strong evidence for the action of diversifying selection on this locus.

The results for P-cadherin also provide evidence for positive selection. Once again, model 2 does not show a category under positive selection, but model 3 does, with approximately 4% of the sites showing an average ω of 2.6. Model 8 gives essentially the same results. The empirical Bayes approach for both model 3 and model 8 produced almost identical results, with significant probabilities of positive selection estimated at six sites for model 3, and for five of those same six sites for model 8 (the sixth site barely missed significance). The sites estimated to be under positive selection were codons 57, 61, 70, 89, 339 and 648. Each model shows high probabilities of positive selection at 21 sites. Hence, both models identify the same sites as focal points for positive selection on this molecule. The LRTs were all highly significant, providing statistical support for the estimates of positive selection.

The results for VE-cadherin were mixed, although they did provide some support for positive selection on this locus. Again, model 2 did not have a category under positive selection, whereas model 3 showed about 2.6% of the sites with an average ω of 2.01. Model 8 estimated a ω of approximately 2.65 for 1.2% of the sites. The Bayesian analysis did not show any sites with significant probabilities of positive selection (sites 13, 39 and 152 came close to significance under the discrete (M3) model). Also, the LRT comparing model 8 to model 7 did not show a significant difference, so there was no rigorous statistical support for the action of positive selection on this locus. The LRT comparing model 3 to model 0 was highly significant, demonstrating significant variation in ω among sites. Hence, it is apparent that patterns of selection differ among codons at this locus, but evidence for positive selection (in which ω significantly exceeds one) is relatively weak.

The analysis of H-cadherin showed no evidence of diversifying selection. None of the three selection models (M2, M3 or M8) estimated any sites to be under positive selection.

Overall, two out of the three cadherins involved in maternal–foetal interactions showed strong evidence for positive selection, in spite of the small sample size of sequences available for analysis. The third locus (VE-cadherin) also showed evidence of positive selection, although statistical support was relatively weak. Given

Table 2. Maximum likelihood estimates of parameters from PAML analyses of cadherins.
(Abbreviations are as follows: Data = dataset. M = ML model; np = number of parameters in model; ω_0, ω_1, ω_2 = d*N*/d*S* category for sites under negative, neutral and positive selection, respectively; p_0, p_1, p_2 = proportion of sites in first category, etc.; *p* and *q* = parameters of the beta distribution (continuous distribution approximation); log*L* = log likelihood of the estimated distribution under a particular model; 2dL = two times the difference in log likelihoods between specific models; *P* = *P* value for the specific comparisons (e.g. M0 versus M3), assuming a chi-square distribution.)

data	M	np	ω_0	p_0	ω_1	p_1	ω_2	p_2	*p*	*q*	log*L*	2dL	*P*
Ecad	M0	7	0.187	1							−7421.03		
Ecad	M1	7	0	0.672	1	0.328					−7341.76		
Ecad	M2	9	0	0	1	0.193	0.057	0.807			−7320.62		
Ecad	M3	11	0.065	0.828	1.151	0.172	1.222	0			−7320.27	201.52	$p<0.01$
Ecad	M7	8							0.144	0.484	−7323.33		
Ecad	M8	10		0.992			89.35	0.008	0.163	0.578	−7314.9	16.86	$p<0.01$
Pcad	M0	9	0.134	1							−5220.66		
Pcad	M1	9	0	0.731	1	0.269					−5218.62		
Pcad	M2	11	0	0	1	0.089	0.077	0.911			−5197.62		
Pcad	M3	13	0	0.031	0.102	0.931	2.629	0.037			−5194.99	51.34	$p<0.01$
Pcad	M7	10							0.24	1.223	−5201.19		
Pcad	M8	12		0.964			2.661	0.036	10.998	98.994	−5194.99	12.4	$p<0.01$
VEcad	M0	9	0.174	1							−6932.12		
VEcad	M1	9	0	0.594	1	0.406					−6888.83		
VEcad	M2	11	0	0.456	0.12	1	0.424	0.25			−6839.86		
VEcad	M3	13	0.028	0.626	0.469	0.348	2.01	0.026			−6839.47	185.3	$p<0.01$
VEcad	M7	10							0.25	0.873	−6840.74		
VEcad	M8	12		0.988			2.653	0.012	0.285	1.087	−6839.52	2.44	N.S.
Hcad	M0	9	0.083	1							−5730.47		
Hcad	M1	9	0	0.691	1	0.309					−5784.28		
Hcad	M2	11	0	0.522	1	0.014	0.183	0.464			−5704.58		
Hcad	M3	13	0.023	0.103	0.023	0.623	0.293	0.274			−5704.38		
Hcad	M7	10							0.372	3.288	−5704.47		
Hcad	M8	12		0.736			0.297	0.264	2.631	99	−5704.39		

the small number of sequences analysed, this could be a problem of insufficient power. Future analyses employing a larger number of sequences may provide stronger statistical support for the action of positive selection on this gene. H-cadherin, which is not known to be involved in maternal–foetal interactions, did not show any evidence of positive selection.

Our results indicate that the cadherin genes involved in maternal–foetal interactions have been subject to diversifying selection. In turn, this is consistent with the hypothesis that these cadherin genes have been subject to green-beard-effect mutations over the course of evolutionary history, leading to antagonistic coevolution with suppressing elements from the parliament of genes. The main alternative hypotheses to this one are twofold. First, cadherins may have been subject to positive selection in the context of maternal–foetal interactions not mediated by green-beard effects. For example, Crespi & Semeniuk (2004) describe the evidence for positive selection on other genes associated with placentation, and Lecuit *et al.* (2004) show how a bacterial pathogen can cross the human maternofoetal barrier via a process mediated in part by E-cadherin. Second, given that some of the cadherins involved in maternal–foetal interactions are also known to be expressed in other tissues, they may have been subject to positive selection in other contexts, such as regulation of development (Noonan *et al.* 2003; Wheelock & Johnson 2003) or the evolution of vulnerability to pre-eclampsia (Pang & Xing 2003) or cancer (Ilyas 2000; Hirohashi & Kanai 2003).

The evolutionary and clinical implications of the green-beard hypothesis and its alternatives should strongly motivate further research on cadherin molecular evolution. The green-beard hypothesis can be tested further, and differentiated from the alternatives, via: (i) studies of the expected molecular effects of variant amino acids at the sites inferred to have been subject to positive selection, (ii) genomic studies that characterize the within-population haplotype variability in these genes, (iii) tests for green-beard effects in pregnancies of mothers heterozygous for placentally expressed cadherins and (iv) analysis of the molecular evolution of genes that modify cadherin activity during trophoblast and placental development.

NOTE ADDED IN PROOF

The recommended selection models in PAML have been changed since this paper was accepted (see documentation for version 3.14 at the PAML website http://abacus.gene.ucl.ac.uk/software/paml.html). We have implemented the modified versions of Model 2 and Model 8, as recommended in the PAML documentation. The results remain qualitatively the same as those presented here.

We thank David Haig, David Queller, Mitsu Ikura, Robert Trivers, an anonymous reviewer and the SFU Fab-Laboratory for helpful comments and NSERC for financial support.

REFERENCES

Alexander, R. D. & Borgia, G. 1978 Group selection, altruism, and the levels of organization of life. *Annu. Rev. Ecol. Syst.* **9**, 449–474.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

Anisomova, M., Bielawski, J. P. & Yang, Z. 2001 Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol. Biol. Evol.* **18**, 1585–1592.

Anisomova, M., Bielawski, J. P. & Yang, Z. 2002 Accuracy and power of the Bayes prediction of amino acid sites under positive selection. *Mol. Biol. Evol.* **19**, 950–958.

Blaschuk, O. W., Sullivan, R., David, S. & Pouliout, Y. 1990 Identification of a cadherin cell adhesion recognition sequence. *Dev. Biol.* **139**, 227–229.

Campbell, S., Swann, H. R., Seif, M. W., Kimber, S. J. & Aplin, J. D. 1995 Cell adhesion molecules on the oocyte and preimplantation human embryo. *Hum. Reprod.* **10**, 1571–1578.

Crespi, B. & Semeniuk, C. 2004 Parent–offspring conflict in the evolution of vertebrate reproductive mode. *Am. Nat.* **163**, 635–653.

Crespi, B. & Springer, S. 2003 Social slime molds meet their match. *Science* **299**, 56–57.

Dawkins, R. 1976 *The selfish gene*. Oxford: Oxford University Press.

Edelman, G. M., Murray, B. A., Mege, R. M., Cunningham, B. A. & Gallin, W. J. 1987 Cellular expression of liver and neural cell adhesion molecules after transfection with their cDNAs results in specific cell–cell binding. *Proc. Natl Acad. Sci. USA* **84**, 8502–8506.

Floridon, C., Nielsen, O., Holund, B., Sunde, L., Westergaard, J. G., Thomsen, S. G. & Teisner, B. 2000 Localization of E-cadherin in villous, extravillous and vascular trophoblasts during intrauterine, ectopic and molar pregnancy. *Mol. Hum. Reprod.* **6**, 943–950.

Friedlander, D. R., Mege, R. M., Cunningham, B. A. & Edelman, G. M. 1989 Cell sorting-out is modulated by both the specificity and amount of different cell adhesion molecules (CAMs) expressed on cell surfaces. *Proc. Natl Acad. Sci. USA* **86**, 7043–7047.

Gallin, W. J. 1998 Evolution of the 'classical' cadherin family of cell adhesion molecules in vertebrates. *Mol. Biol. Evol.* **15**, 1099–1107.

Gallin, W. J., Edelman, G. M. & Cunningham, B. A. 1983 Characterization of L-CAM, a major cell adhesion molecule from embryonic liver cells. *Proc. Natl Acad. Sci. USA* **80**, 1038–1042.

Getsios, S., Chen, G. T. & MacCalman, C. D. 2000 Regulation of beta-catenin mRNA and protein levels in human villous cytotrophoblasts undergoing aggregation and fusion in vitro: correlation with E-cadherin expression. *J. Reprod. Fertil.* **119**, 59–68.

Guilford, T. 1988 The evolution of conspicuous coloration. In *Mimicry and the evolutionary process* (ed. L. P. Brower), pp. 7–21. Chicago: University of Chicago Press.

Haig, D. 1993 Genetic conflicts in human pregnancy. *Q. Rev. Biol.* **68**, 495–532.

Haig, D. 1996 Gestational drive and the green-bearded placenta. *Proc. Natl Acad. Sci. USA* **93**, 6547–6551.

Haig, D. 1997 Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proc. R. Soc. B* **264**, 1657–1662.

Haig, D. 1999 Altercation of generations: genetic conflicts of pregnancy. *Am. J. Reprod. Immunol.* **35**, 226–232.

Haig, D. 2000 The kinship theory of genomic imprinting. *Annu. Rev. Ecol. Syst.* **31**, 9–32.

Haig, D. & Graham, C. 1991 Genomic imprinting and the strange case of the insulin-like growth factor-II receptor. *Cell* **64**, 1045–1046.

Hamilton, W. D. 1964 The genetical theory of social behavior I & II. *J. Theor. Biol.* **7**, 1–52.

Hurst, L., Atlan, A. & Bengtsson, B. 1996 Genetic conflicts. *Q. Rev. Biol.* **71**, 317–364.

Hirohashi, S. & Kanai, Y. 2003 Cell adhesion system and human cancer morphogenesis. *Cancer Sci.* **94**, 575–581.

Hyafil, F., Morello, D., Babinet, C. & Jacob, F. 1980 A cell surface glycoprotein involved in the compaction of embryonal carcinoma cells and cleavage stage embryos. *Cell* **21**, 927–934.

Ilyas, M. 2000 Adhesion molecule expression in breast cancer: the phoenix in tumour metastasis? *J. Pathol.* **190**, 3–5.

Keller, L. & Ross, K. G. 1998 Selfish genes: a green beard in the red fire ant. *Nature* **394**, 573–575.

King, N., Hittinger, C. T. & Carroll, S. B. 2003 Evolution of key cell signaling and adhesion protein families predates animal origins. *Science* **301**, 361–363.

Klymkowsky, M. W. & Parr, B. 1995 The body language of cells: the intimate connection between cell adhesion and behavior. *Cell* **83**, 5–8.

Lecuit, M., Nelson, D. M., Smith, S. D., Khun, H., Huerre, M., Vacher-Lavenu, M. C., Gordon, J. I. & Cossart, P. 2004 Targeting and crossing of the human maternofetal barrier by *Listeria monocytogenes*: role of internalin interaction with trophoblast E-cadherin. *Proc. Natl Acad. Sci.* **101**, 6152–6157.

Leigh, E. G. 1971 *Adaptation and diversity.* San Francisco: Freeman.

MacCalman, C. D., Furth, E. E., Omigbodun, A., Bronner, M., Coutifaris, C. & Strauss, J. F. 1996 Regulated expression of cadherin-11 in human epithelial cells: a role for cadherin-11 in trophoblast–endometrium interactions? *Dev. Dyn.* **206**, 201–211.

Nagar, B., Overduin, M., Ikura, M. & Rini, J. M. 1986 Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature* **380**, 360–364.

Nielsen, R. & Yang, Z. 1998 Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1. envelope gene. *Genetics* **148**, 929–936.

Nollet, F., Kools, P. & van Roy, F. 2000 Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J. Mol. Biol.* **299**, 551–572.

Noonan, J. P., Li, J., Nguyen, L., Caoile, C., Dickson, M., Grimwood, J., Schmutz, J., Feldman, M. W. & Myers, R. M. 2003 Extensive linkage disequilibrium, a common 16.7-kilobase deletion, and evidence of balancing selection in the human protocadherin alpha cluster. *Am. J. Hum. Genet.* **72**, 621–635.

Nose, A. & Takeichi, M. 1986 A novel cadherin cell adhesion molecule: its expression patterns associated with implantation and organogenesis of mouse embryos. *J. Cell Biol.* **10**, 2649–2658.

Overduin, M., Harvey, T. S., Bagby, S., Tong, I. T., Yau, P., Takeichi, M. & Ikura, M. 1995 Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. *Science* **267**, 386–389.

Page, R. D. M. 1996 TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358.

Pang, Z. J. & Xing, F. Q. 2003 Expression profile of trophoblast invasion-associated genes in the pre-eclamptic placenta. *Br. J. Biomed. Sci.* **60**, 97–101.

Paria, B. C., Zhao, X., Das, S. K., Dey, S. K. & Yoshinaga, K. 1999 Zonula occludens-1 and E-cadherin are coordinately expressed in the mouse uterus with the initiation of implantation and decidualization. *Dev. Biol.* **208**, 488–501.

Pertz, O., Bozic, D., Koch, A. W., Fauser, C., Brancaccio, A. & Engel, J. 1999 A new crystal structure, Ca^{2+} dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. *EMBO J.* **18**, 1738–1747.

Queller, D. C. 1984 Kin selection and frequency dependence: a game-theoretic approach. *Biol. J. Linn. Soc.* **23**, 133–143.

Queller, D. C., Ponte, E., Bozzaro, S. & Strassmann, J. E. 2003 Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. *Science* **299**, 105–106.

Ridley, M. & Grafen, A. 1981 Are green beard genes outlaws? *Anim. Behav.* **29**, 954–955.

Schuh, R., Vestweber, D., Riede, L., Ringwald, M., Rosenberg, U. B., Jackle, H. & Kemler, R. 1986 Molecular cloning of the mouse cell adhesion molecule uvomorulin: cDNA contains a B1-related sequence. *Proc. Natl Acad. Sci. USA* **83**, 1364–1368.

Shapiro, L. *et al.* 1995*a* Structural basis of cell–cell adhesion by cadherins. *Nature* **374**, 327–337.

Shapiro, L., Kwong, P. D., Fannon, A. M., Colman, D. R. & Hendrickson, W. A. 1995*b* Considerations on the folding topology and evolutionary origin of cadherin domains. *Proc. Natl Acad. Sci. USA* **92**, 6793–6797.

Shih, I. M., Hsu, M. Y., Oldt, R. J., Herlyn, M., Gearhart, J. D. & Kurman, R. J. 2002 The role of E-cadherin in the motility and invasion of implantation site intermediate trophoblast. *Placenta* **23**, 706–715.

Springer, M. S., Stanhope, M. J., Madsen, O. & de Jong, W. W. 2004 Molecules consolidate the placental mammal tree. *TREE* **19**, 430–438.

Swofford, D. L. 2002 *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sunderland, Massachusetts: Sinauer Associates.

Takeichi, M. 1991 Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* **251**, 1451–1455.

Tomschy, A., Fauser, C., Landwehr, R. & Engel, J. 1996 Homophilic adhesion of E-cadherin occurs by a co-operative two-step interaction of N-terminal domains. *EMBO J.* **15**, 3507–3514.

Wernersson, R. & Pederson, A. G. 2003 RevTrans—constructing alignments of coding DNA from aligned amino acids sequences. *Nucleic Acids Res.* **31**, 3537–3539.

Wheelock, M. J. & Johnson, K. R. 2003 Cadherins as modulators of cellular phenotype. *Annu. Rev. Cell Dev. Biol.* **19**, 207–235.

Yang, Z. 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS* **13**, 555–556.

Yang, Z. 2001 Adaptive molecular evolution. In *Handbook of statistical genetics* (ed. D. J. Balding). New York: Wiley.

Yang, Z. & Bielawski, J. P. 2000 Statistical methods to detect molecular adaptation. *Trends Ecol. Evol.* **15**, 495–503.

Yang, Z., Nielsen, R., Goldman, N. & Krabbe Pedersen, A. M. 2000 Codon-substitution models for heterogenous selection pressure at amino acid sites. *Genetics* **155**, 431–439.

Zhou, Y., Fisher, S. J., Janatpour, M., Genbacev, O., Dejana, E., Wheelock, M. & Damsky, C. H. 1997 Human cytotrophoblasts adopt a vascular phenotype as they differentiate A strategy for successful endovascular invasion? *J. Clin. Invest.* **99**, 2139–2151.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.

The supplementary Electronic Appendix is available at http://dx.doi.org/10.1098/rspb.2004.2890 or via http://www.journals. royalsoc.ac.uk.