Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking

Jonathan Benjamin[1], Lin Li[2], Chavis Patterson[2], Benjamin D. Greenberg[1], Dennis L. Murphy[1] & Dean H. Hamer[2]

Twin and adoption studies suggest that 30 to 60% of the variance in many personality traits is due to inherited factors. However, there is little knowledge of the number or identity of the responsible genes, how they differ between individuals, or how their gene products interact with the developing brain and with environmental and experiential factors to generate the complex blend of attitudes and actions that comprise human temperament[1]. In the accompanying paper, Ebstein et al.[2] have found a population association between a long allele of polymorphic exon III repeat sequence of the D4 dopamine receptor gene (*D4DR*) and the normal personality trait of Novelty Seeking. The possibility of a causal relationship between *D4DR* and Novelty Seeking is further supported by studies showing that the number of exon III repeats can affect the binding of ligands to the receptor[3,4]; that *D4DR* is expressed in limbic areas involved in cognition and emotion[5,6]; that dopamine mediates exploratory behaviour in experimental animals[7–12]; that the rewarding effects of amphetamines and cocaine are related to dopamine release[13]; and that Novelty Seeking is low in dopamine-deficient patients with Parkinson's disease[14]. We investigated the relationship between *D4DR* exon III sequence variants and personality test scores in a population of 315 mostly male siblings, other family members and individuals from the United States. The association between long alleles of exon III and personality traits related to Novelty Seeking was confirmed. Moreover, family studies showed that this association is the result of genetic transmission rather than of population stratification.

Ebstein et al.[2] measured personality traits using Cloninger's tridimensional personality questionnaire (TPQ), which is based on a biological model in which three independent dimensions of temperament are attributed to genetically and neurochemically distinct pathways[12,15]. It was hypothesized that Novelty Seeking, which is characterized by "exhilaration or excitement in response to novel stimuli", is mediated by dopamine neurotransmission[15]. Our studies utilized the NEO personality inventory (NEO-PI-R), which is based on the five-factor model of personality structure[16]. This questionnaire was chosen because it has high retest reliability and longitudinal stability, its factor structure has been validated in a variety of populations and cultures, and there is a good correlation between self reports and observer ratings[17,18]. Although the NEO-PI-R does not include Novelty

Seeking as a specific factor, it contains multiple items that are clearly related to questions from the TPQ-Novelty Seeking scale; for example, "I have sometimes done things just for kicks or thrills" versus "I often try new things just for fun or thrills", and "I think things through before coming to a decision" versus "I like to think about things for a long time before I make a decision." Recent empirical studies have demonstrated an overall correlation of approximately 70% between the NEO-PI-R factors and the TPQ-Novelty Seeking scale, largely attributable to positive loading on Extraversion and negative loading on Conscientiousness (C.R. Cloninger, personal communication)[17]. Based on these analyses and the prior results of Ebstein et al.[2], we hypothesized that long alleles of *D4DR* exon III would be positively associated with NEO-PI-R-Extraversion, negatively associated with NEO-PI-R-Conscientiousness, and positively associated with TPQ-Novelty Seeking scores estimated from our data by appropriately weighted equations[18].

The lengths of the *D4DR* exon III-repeat sequences were determined for 315 people for whom NEO-PI-R results were available. The genotypes were divided into two groups: those containing only the short (S) *D4DR* alleles with 2 to 5 exon III repeats ($n = 217$), and those containing one or two copies of the long (L) *D4DR* alleles with 6 to 8 exon III repeats ($n = 98$). Similar results were obtained when genotypes were stratified according to Ebstein et al.[2] by the presence or absence of allele 7 (the most common long allele) or by 4/4 versus 4/7 genotype (the most common S and L genotypes), when the number of long alleles was taken into account (S/S versus S/L versus L/L genotype), when the lengths of the two alleles were added together and treated as a quantitative variable, or when only the length of the longest allele was considered (see Methods). Although exon III is polymorphic with regard to sequence as well as length[19], it is not known whether sequence variations affect the biological activity of the receptor, and we did not sequence the exon III region in this study. It is conceivable that the observed associations for different sized *D4DR* exon III alleles are due to linkage disequilibrium with a particular sequence variant rather than length *per se*.

Table 1*a* shows a one-way ANOVA of the five major NEO-PI-R personality factors with simplified *D4DR* genotype as the independent factor. Because of the diverse nature of the study participants, the personality scores were statistically corrected for age, sex, ethnicity and sexual orientation (see Methods). As hypothesized, the scores for Extraversion were significantly higher ($P = 0.001$) and the scores for Conscientiousness were significantly lower ($P = 0.03$) in L than in S subjects. The other three NEO-PI-R personality factors were not significantly associated with *D4DR* genotype.

The NEO-PI-R is based on a hierarchal model in which each of the five major personality factors is subdivided into 6 facets[16]. An analysis of the individual facets that comprise Extraversion and Conscientiousness is shown in Table 1*b*. The L subjects scored higher than S subjects on three out of the six facets of Extraversion: Warmth ($P = 0.005$), Excitement-Seeking ($P = 0.003$) and Positive Emotions ($P = 0.0009$). For the Conscientiousness factor, *D4DR* genotype was significantly associated with the single facet of Delibera-

[1]*Laboratory of Clinical Science, National Institute of Mental Health and* [2]*Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892*

Correspondence should be addressed to D.H.

letters

Table 1 Population associations between *D4DR* genotype and personality traits

| | | Mean T score (SD)[a] | | | | |
		S group (n = 217)	L group (n = 98)	F[b]	ΔD4DR[c]	P value
a, NEO-PI-R Factors						
N	Neuroticism	53.1(9.1)	52.3(9.2)	0.5	ns	ns
E	Extraversion	53.4(9.7)	57.3(10.3)	10.7	3.9	.0012
O	Openness	58.5(9.7)	59.7(11.1)	0.9	ns	ns
A	Agreeableness	46.1(9.5)	48.3(9.9)	3.6	ns	ns
C	Conscientiousness	45.9(10.3)	43.2(10.7)	4.7	−2.7	.0311
b, NEO-PI-R Extraversion and Conscientiousness Facets						
E1	Warmth	49.3(10.1)	52.6(8.9)	7.9	3.4	.0051
E2	Gregariousness	54.4(10.2)	55.9(9.4)	1.6	ns	ns
E3	Assertiveness	51.5(9.3)	51.6(10.8)	0.0	ns	ns
E4	Activity	54.0(9.0)	53.3(11.0)	0.4	ns	ns
E5	Excitement Seeking	54.2(8.6)	57.4(8.6)	9.0	3.1	.0030
E6	Positive Emotions	52.6(9.7)	56.5(9.7)	11.2	3.9	.0009
C1	Competence	50.8(9.5)	50.5(10.7)	0.1	ns	ns
C2	Order	46.3(9.5)	44.9(10.1)	1.2	ns	ns
C3	Dutifulness	46.8(9.9)	44.8(9.6)	2.8	ns	ns
C4	Achievement Striving	48.8(9.6)	47.4(10.1)	1.4	ns	ns
C5	Self Discipline	41.7(9.5)	42.6(11.0)	0.5	ns	ns
C6	Deliberation	47.6(9.9)	44.4(9.2)	7.1	−3.2	.0080
c, TPQ-Novelty Seeking Scale and Subscales[d]						
NS	Novelty Seeking	55.1(7.9)	58.1(7.0)	10.2	3.0	.0016
NS1	Expl. Excitability	54.3(6.6)	57.0(6.4)	11.4	2.7	.0008
NS2	Impulsiveness	53.6(4.5)	54.6(4.5)	3.6	ns	ns
NS3	Extravagance	52.0(5.9)	54.2(5.3)	10.5	2.3	.0013
NS4	Disorderliness	54.1(6.1)	56.1(5.6)	7.9	2.0	.0051

[a]Personality test scores are given as T scores, which are standardized to have a mean (±SD) of 50 (±10) in the NEO-PI-R normative sample.
[b]F value by one-way ANOVA.
[c]ΔD4DR = (mean T score for L group) - (mean T score for S group).
[d]Estimated from NEO-PI-R factor scores by weighted equations.

jects were genetically related. George & Elston[20] have described a test for associations between genetic markers and quantitative traits in pedigree data that corrects for the statistical dependence among family members[21,22]. Table 2 shows that NEO-PI-R- Extraversion, NEO-PI-R-Conscientiousness and estimated TPQ- Novelty Seeking scores were significantly associated to *D4DR* genotype as determined by this likelihood ratio test. The effect sizes and significance levels were similar to those obtained by ANOVA of the same data.

A population association between a genetic polymorphism and a phenotype can arise either because of a direct relationship between the gene and the trait (genetic transmission) or because subpopulations that have different frequencies of the polymorphism also happen to differ in average phenotype (population stratification). These alternatives can be distinguished by examining associations within pedigrees rather than across pedigrees[23]. Table 3 shows an analysis of all 60 sib-pairs in our sample in which one sib had an L genotype and one sib had an S genotype. A paired-samples t-test showed that the L siblings scored significantly higher on NEO-PI-R- Extraversion, lower on NEO-PI-R-Conscientiousness, and higher on estimated TPQ-Novelty Seeking than the S siblings from the same family. The pairwise differences between S and L siblings were not significantly different from the average differences between all S and L individuals (Table 1) and were statistically significant after two different conservative corrections for the nonindependence between sib-pairs from the same pedigree (see Methods). These results indicate that the association between the *D4DR* exon III polymorphism and measures of Novelty Seeking is due to genetic transmission rather than population stratification.

In considering the relationship between a genetic polymorphism and a quantitative trait, it is important

tion, and as expected the scores were lower in L than in S subjects (P = 0.008).

TPQ-Novelty seeking scores were estimated from the NEO-PI-R test results by weighted equations[18]. Table 1c shows that long alleles of D4DR exon III were positively associated with the overall dimension of Novelty Seeking (P = 0.002) and with the Novelty Seeking subscales of Exploratory Excitability (P = 0.0008), Extravagance (P = 0.001) and Disorderliness (P = 0.005). The magnitudes and significances of the associations to the estimated TPQ-Novelty Seeking scores were similar to those observed for the directly measured NEO-PI-R scores.

In the above analyses, each individual was treated as an independent datum even though many of the sub-

Table 2 Across pedigrees association test

		ΔD4DR[a]	−2LnL[b]	P[c]
E	Extraversion[d]	4.2	11.6	0.0006
C	Conscientiousness[d]	−2.7	4.6	0.0318
NS	Novelty Seeking[e]	3.1	10.6	0.0011

[a]ΔD4DR = maximum likelihood estimate of (T score for L group) – (T score for S group).
[b]−2LnL = −2 × (log(likelihood of data without ΔD4DR) – log(likelihood of data with ΔD4DR)).
[c]P was calculated by taking −2LnL to be distributed as a chi-squared statistic at one degree of freedom.
[d]NEO-PI-R factor score.
[e]TPQ score estimated from NEO-PI-R factor scores by a weighted equation.

Table 3 Within pedigrees association test[a]

		ΔD4DR[b]	t[c]	P
E	Extraversion[4]	6.8	2.9	0.008
C	Conscientiousness[4]	−4.7	2.3	0.032
NS	Novelty Seeking[5]	4.6	2.8	0.010

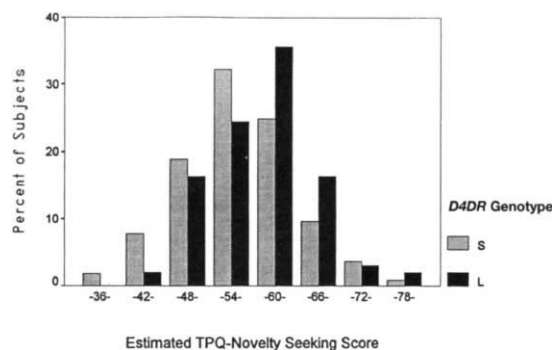[a]n = 31 pedigrees. Ixs = 60 sib-pairs. See Methods for details.
[b]ΔD4DR = mean (T score for L sib - T score for S sib).
[c]t-score conservatively corrected for nonindependence of sib-pairs from a single family.
[d]NEO-PI-R factor score.
[e]TPQ score estimated from NEO-PI-R factor scores by a weighted equation.

Fig. 1 Distributions of estimated Novelty Seeking scores. The X axis shows the estimated TPQ-Novelty Seeking scores separated into 8 groups with the indicated median T scores. The Y axis shows the distribution, in percent, of subjects with short *D4DR* exon III alleles (group S, n = 217, hatched bars) in each of the eight groups and the distribution, in percent, of subjects with long *D4DR* exon III alleles (group L, n = 98, solid bars) in each of the eight groups.

to consider the distribution as well as the mean of the measured phenotype. Fig. 1 shows histograms of estimated TPQ-Novelty Seeking scores in the two *D4DR* genotype classes. Although the mean score for the L subjects is greater than for the S subjects by 0.4 standard deviations (a moderate effect size)[24], the distributions are highly overlapping and *D4DR* accounts for only 3 to 4% of total variance. The broad heritability of Novelty Seeking has been estimated to be 41% from twin studies[25], and in our families there was a correlation of 0.23 for estimated TPQ–Novelty Seeking scores in siblings. Thus, *D4DR* accounts for roughly 10% of the genetic variance, as might be anticipated if there are 10 or so genes for this complex, normally distributed trait. These results indicate that Novelty Seeking is partially but not completely mediated by genes, and that the *D4DR* polymorphism accounts for some but not all of the genetic effects.

These results confirm and extend those of Ebstein *et al.*[2] of an association between long alleles of *D4DR* exon III and Novelty Seeking. The convergence of our results is especially striking considering that we used different personality questionnaires and ethnically disparate populations (Ashkenazi and Sephardic Jews from Israel versus a mixed group of predominantly non-Jewish Caucasians as well as Hispanics, Asians and African Americans from the United States). Association tests of the sort described here provide a powerful method for detecting the subtle effects of genes on multifactorial quantitative traits, especially when biologically reasonable candidate loci are available[26,27]. A similar approach may be useful for detecting genes that influence abnormal psychological processes and health risk-related behaviors such as tobacco smoking and excess alcohol consumption.

Methods

The subjects were participants in two NIH protocols: a National Institute of Mental Health project on "Mapping Personality Traits to Chromosomes" and a National Cancer Institute project on "Genetic Factors and Interrelationships for Sexual Orientation, HIV Progression, Alcoholism and Psychological Traits, and Histocompatibility Antigens." Because both protocols began as studies of the X chromosome, they focus on male siblings and their family members. Subjects were recruited from the NIH campus and clinics, local universities, and local and national homophile organizations through advertisements and word of mouth. The NIMH protocol subjects were paid a nominal sum for their participation.

There were 315 subjects of whom 95% were male and 5% were female. The ethnic composition was 92% white non-Hispanic, 4.1% Asian/Pacific Islander, 2.9% Hispanic/Latino, 0.6%

African American/Black, and 0.3% Native American/Alaskan. The average age was 32.4 ± 10.8 years (range 18–71) and the average educational level was 16.7 ± 2.7 years (range 12–20 years). The sample was approximately evenly split with regard to heterosexual and homosexual orientation with an average Kinsey scale rating of 3.0 ± 2.8 (range 0 to 6). The family structure of the sample was 291 siblings from 131 families, 7 parents, and 17 unrelated individuals.

After completing informed consent, subjects were given a questionnaire or structured interview that covered demographics, brief medical, genetic and psychiatric history, tobacco, alcohol and drug use, and sexual and gender-related behaviors and attitudes[28]. They were then asked to complete the NEO-PI-R form S, a self-report pencil and paper test consisting of 240 items each with five possible answers (strongly disagree to strongly agree)[16]. Participants also completed the Cattell 16PF inventory, but because this questionnaire has not been directly compared to the TPQ those data are not considered here.

Subjects donated a blood sample from which DNA was extracted by a commercial service (Genetic Design, Greensboro, NC). The *D4DR* exon III polymorphism was assayed by PCR using the D4-3 and D4-42 primers and action conditions as described[19] combined with a "hot start". The amplified fragments were separated by 3.5% Metaphor agarose gel electrophoresis and allele numbers were assigned according to the number of repeats[3]. The allele frequencies in our population of 630 chromosomes were: allele 2, 8.9%; allele 3, 4.4%; allele 4, 67.3%; allele 5, 1.7%; allele 6, 0.2%; allele 7, 16.7%; allele 8, 0.8%.

Based on Ebstein *et al.*'s finding of a difference in Novelty Seeking between individuals with 4- and 7- exon III repeats[2], together with the evidence that the 4 and 7 repeat proteins have different ligand-binding properties[3,4], alleles were assigned to two bins at the midpoint between these variants (s = alleles 2–5, l = alleles 6–8), and genotypes were grouped as S = s/s, L = s/l or l/l. The following statistical results for estimated TPQ-Novelty Seeking scores were obtained when different genotype classifications were employed: For S vs L genotype, F = 10.2, P = 0.002; for s/s vs s/l vs l/l genotype, F = 5.5, P = 0.004; for 4/4 vs 4/7 genotype, F = 11.3, P = 0.001; for genotypes with allele 7 vs genotypes without allele 7, F = 8.2, P = 0.005; for sum of allele lengths as a quantitative variable, F = 4.4, P = 0.037; for length of the longest allele, F = 9.4, P = 0.002.

All personality test scores are reported as T scores, which are standardized to have a mean (± standard deviation) of 50 (±10) in the NEO-PI-R normative population. Factor scores were calculated from the scoring weights provided in the NEO-PI-R manual[16]. TPQ-Novelty Seeking scores were estimated as a weighted combination of NEO-PI-R factor scores, with weights determined from a correlation matrix obtained by simultaneously administering the TPQ and the NEO-PI-R to a population of college students[18]. The estimated TPQ-Novelty Seeking scores are reported as T-scores rather than as the raw scores in Ebstein *et al.*[2]. The equation for the TPQ-Novelty Seeking scale was NS = $-0.09 \times N + 0.32 \times E + 0.17 \times O - 0.1 \times A - 0.6 \times C + 35$ where N, E, O, A, and C are NEO-PI-R factor scores. Similar results were obtained when an equation based on an independent dataset provided by C.R. Cloninger was employed. All personality test scores were statistically corrected by regression against age (standardized regression coefficient for estimated TPQ-Novelty seeking = β = -0.26, $P < 0.0001$), sex (β = 0.02, $P > 0.2$), ethnicity (β = 0.01, $P > 0.2$) and sexual

orientation (average Kinsey score, $\beta = -0.07$, $P > 0.2$). The significant negative correlation between age and estimated TPQ-Novelty Seeking was expected from normative data[12,29]. However, the association between estimated TPQ-Novelty Seeking scores and *D4DR* genotype remained significant even when uncorrected scores were used.

The strength and significance of associations between *D4DR* genotype and personality test scores were determined by one-way ANOVA (SPSS statistical software). Because the experimental design and analysis were based on the prior results of Ebstein *et al.*[2] significance levels are reported as direct probability values. The across pedigrees association test was implemented with the ASSOC program of the S.A.G.E package[22]. The within pedigrees association test (Table 3) examined all N families in which there was one sib with an L genotype and one sib with an S genotype, each of which formed lxs pairs (l is the number of L sibs, s is the number of S sibs); for example, a family with two S sibs and one L sib would give two S-L pairs and a family with one S sib and three L sibs would give three S-L sib-pairs. A paired-samples t-test was used to determine the mean T score difference between L and S sibs ($\Delta D4DR$) and the associated t-statistic for this population of $\sum_{l \times s}$ sib-pairs.

Because the sib-pairs from a single family are not statistically independent, the resulting t-score was scaled down according to the equation $t^* = t \times (N/\sum_{l \times s})^{1/2}$, and the significance level was conservatively estimated by evaluating t^* as a t-statistic at N-1 degrees of freedom (David Fulker, personal communication). Similar results were obtained for estimated TPQ-Novelty Seeking scores when only the first sib-pair from each family was analysed ($t = 2.2$, D.F. $= 30$, $P = 0.037$).

1. Bouchard, T.J.J. Genes, environment, and personality. *Science* **264**, 1700–1701 (1994).
2. Ebstein, R.P. *et al.* Dopamine D4 receptor (*D4DR*) exon III polymorphism associated with the human personality trait of Novelty Seeking. *Nature Genet.* **12**, 78–80 (1996).
3. Van Tol, H.H.M. *et al.* Multiple dopamine D4 receptor variants in the human population. *Nature* **358**, 149–152 (1992).
4. Asghari, V. *et al.* Dopamine D4 receptor repeat: analysis of different native and mutant forms of the human and rat genes. *Mol. Pharmacol.* **46**, 364–373 (1994).
5. Van Tol, H.H.M. *et al.* Cloning of the gene for a human dopamine D4 receptor with high affinity for the antipsychotic clozapine. *Nature* **350**, 610–614 (1991).
6. Meador-Woodruff, J.H. *et al.* Dopamine receptor gene expression in the human medial temporal lobe. *Neuropsychopharmacology* **10**, 239–248 (1994).
7. Stellar, J.R. & Stellar, E. in *The Neurobiology of Motivation and Reward* (Springer-Verlag, New York, 1985).
8. Kalivas, P.W. & Nemeroff, C.B. in *The Mesocorticolimbic Dopamine System* 537 (Annals of the New York Academy of Sciences, New York, 1988).
9. Hooks, M.S. & Kalivas, P.W. Involvement of dopamine and excitatory amino acid transmission in novelty-induced motor activity. *J. Pharmacol. Exp. Ther.* **269**, 976–988 (1994).
10. Fink, J.S. & Reis, D.J. Genetic variations in midbrain dopamine cell number: parallel with differences in responses to dopaminergic agonists and in naturalistic behaviors mediated by central dopaminergic systems. *Brain Res.* **222**, 335–349 (1981).
11. Fink, J.S. & Smith, G.P. Mesolimbic and mesocortical dopaminergic neurons are necessary for normal exploratory behavior in rats. *Neurosci. Lett.* **17**, 61–65 (1980).
12. Cloninger, C.R., Przybeck, T.R. & Svrakic, D.M. The Tridimensional Personality Questionnaire: U.S. normative data. *Psychol. Rep.* **69**, 1047–1057 (1991).
13. Ritz, M.C. & Kuhar, M.J. Psychostimulant drugs and a dopamine hypothesis regarding addiction: update on recent research. *Biochem. Soc. Symp.* **59**, 51–64 (1993).
14. Menza, M.A., Golbe, L.I., Cody, R.A. & Forman, N.E. Dopamine-related personality traits in Parkinson's disease. *Neurology* **43**, 505–508 (1993).
15. Cloninger, C.R. A systematic method for clinical description and classification of personality variants. A proposal. *Arch. Gen. Psych.* **44**, 573–588 (1987).
16. Costa, P.T.J. & McCrae, R.R. in *Revised NEO Personality Inventory (NEO PI-R) and NEO Five Inventory (NEO-FFI) professional manual* (Psychological Assessment Resources, Odessa, FL, 1992).
17. McCrae, R.R. & Costa, P.T.J. in *Personality in Adulthood* (Guilford, New York, 1990).
18. Costa, P.T.J. & McCrae, R.R. in *The Revised NEO Personality Inventory (NEO PI-R)* (eds Cheek, J. & Donahue, E.M.) (Plenum, New York, In press).
19. Lichter, J.B. *et al.* A hypervariable segment in the human dopamine receptor D4 (DRD4) gene. *Hum. Mol. Genet.* **2**, 767–773 (1993).
20. George, V.T. & Elston, R.C. Testing the association between polymorphic markers and quantitative traits in pedigrees. *Genet. Epidemiol.* **4**, 193–201 (1987).
21. Nick, T.G., George, V., Elston, R.C. & Wilson, A.F. Statistical validity for testing associations between genetic markers and quantitative traits in family data. *Genet. Epidemiol.* **12**, 145–146 (1995).
22. Elston, R.C., George, V.T. & Sorant, A.J.M. in *Trait-marker association program (ASSOC). In S.A.G.E. Users Guide, Release 2.2 (Department of Biometry and Genetics* (LSU Medical Center, New Orleans, 1994).
23. Spielman, R.S., McGinnis, R.E. & Ewens, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52**, 506–516 (1993).
24. Cohen, J. in *Statistical Power Analysis for the Behavioral Sciences* (Erlbaum, Hillsdale, NJ, 1988).
25. Heath, A.C., Cloninger, C.R. & Martin, N.G. Testing a model for the genetic structure of personality: a comparison of the personality systems of Cloninger and Eysenck. *J. Pers. Soc. Psychol.* **66**, 762–775 (1994).
26. Plomin, R., Owen, M.J. & McGuffin, P. The genetic basis of complex human behaviors. *Science* **264**, 1733–1739 (1994).
27. Lander, E.S. & Schork, N.J. Genetic dissection of complex traits. *Science* **265**, 2037–2048 (1994).
28. Hamer, D.H. & Copeland, P. in *The Science of Desire* (Simon & Schuester, New York, 1994).
29. Costa, P.T.J. & McCrae, R.R. in *The Developing Structure of Temperament and Personality from Infancy to Adulthood* (eds. Halverson, C.F., Kohnstamm, G.A. & Martin, R.P.) 139–150 (Lawrence Erlbaum Associates, Hillsdale, NJ, in the press).