Choice of Residential Location: Chance, Family Influences, or Genes?

John B. Whitfield,[1] Gu Zhu,[1] Andrew C. Heath,[2] and Nicholas G. Martin[1]

[1]*Genetic Epidemiology Unit, Queensland Institute of Medical Research, Brisbane, Australia*
[2]*Department of Psychiatry, Washington University, St Louis, United States of America*

The choice of where to live would appear to be determined by a combination of economic constraints and personal preferences. We have tested how far this choice is affected by the continuing effects of the environment shared within families, and genetic variation between people, using data from twin studies conducted in Australia. The addresses provided by study participants were categorized as urban, suburban and nonurban, and data were analyzed in three adult age groups. There were significant effects of both shared environment and genes, and the balance between them was affected by both sex and age. Shared environment accounted for some 50% of variation in the youngest group, but only about 10% in the oldest. As shared environmental effects decreased, additive genetic effects increased. These results have implications for internal migration of people within countries and, over the long term, for gene flow within and between populations. They may also be pertinent to the different prevalences of certain psychiatric diseases between city and country locations. Comparisons between countries with different demography are needed to confirm and further characterize these effects.

It is not possible for everyone to live where they would prefer, as competition for the most desired locations means that price limits the options available. Fortunately, this competition and the resulting restricted choice may be mitigated by variation between people in the locations and lifestyles which they prefer. Some may choose an inner-city location, others a remote one, and for many the intermediate suburban zone offers the best reconciliation of competing benefits. In practice, the reasons for choice of location are probably mixed. They may include access to employment, business, educational, cultural or recreational opportunities; affordability; familiarity with one location or type of location, perhaps as a result of growing up there; or emotional attachment to a place or a lifestyle. Such emotional/lifestyle considerations are probably responsible for much internal migration within countries like Australia, and may be separated from the more economic aspects of lifestyle improvement which affect movement between countries (which we do not consider here).

One approach to understanding preferences for different locations is to study the degree of similarity in the choices made by people who are similar to each other, and by those who differ. This may be done by studying people who are biologically related to each other, and similarities between relatives may be attributed to shared family environments or to shared genetic influences. In principle, this may include studies of adoptees or extended pedigrees, but a powerful and widely used method is to compare the similarity of twins within either monozygotic (MZ) or dizygotic (DZ) pairs. Both types of pair nearly always share their early family environment, social and educational background, and place of residence until around the age of twenty years. The members of a MZ pair are genetically identical (apart from postconceptional mutations and gene rearrangements) but DZ pairs share, on average, only half their nuclear genes. A greater similarity between members of MZ than DZ pairs is taken as evidence for a genetic influence on the characteristic in question.

In this paper, we analyze data on the addresses given by participants in two large twin studies on personality and behavior, conducted across Australia during the 1990s. Postcodes have been used to identify the participants' residential locations, and form the basis of classification into urban, suburban and nonurban zones. This has been used to examine the effects of shared environments and shared genes on choice of where to live.

Subjects and Methods

Subjects were recruited through the Australian Twin Registry for questionnaire or telephone interview studies on personality and behavior. They were studied in two groups, Cohort 1 (born before 1965;

Received 9 November, 2004; accepted 22 November, 2004.

Address for correspondence: Dr J. B. Whitfield, Genetic Epidemiology Unit, Queensland Institute of Medical Research, 300 Herston Rd, Herston, Queensland, Australia. E-mail: John.Whitfield@qimr.edu.au

Table 1

Categorization of Subjects' Addresses by Distance From City Centres

Reference city	Urban	Suburban	Nonurban
Melbourne, Sydney	Less than 10 km	10 km to 39.99 km	40 km and over
Adelaide, Brisbane, Hobart, Newcastle, Perth	Less than 5 km	5 km to 19.99 km	20 km and over
Canberra, Darwin	N/A	Less than 10 km	10 km and over

Note: For each subject, the postcode of their address was extracted and the distance to the nearest state capital or other major urban centre (as listed) was determined. Addresses were categorized as urban, suburban or nonurban using the criteria shown.

Heath et al., 1997) and Cohort 2 (born between 1965 and 1971; Nelson et al., 2004). There is evidence that these groups were similar to the Australian population in their social and educational characteristics. A number of comparisons with the Australian Bureau of Statistics data have been made for Cohort 1 and deviations with respect to education, socioeconomic status and prevalence of psychiatric symptoms are slight (Baker et al., 1996; Heath et al., 1997; Jardine & Martin, 1984; Kendler et al., 1995). Even if some differences to the overall population do exist, the aim of this study was to assess genetic and nongenetic sources of variation in the location of the study participants, and these aims would not be affected by the deviations that are likely to be present.

Zygosity was determined from responses to questions about physical similarity and the inability of others to tell them apart, supplemented by blood group information and, in the case of DZ pairs included in subsequent linkage analysis, genome-wide microsatellite genotyping.

Date of birth was used to divide the participants into three age groups; those born before 1951, those born between 1951 and 1964 (both from Cohort 1), and those born between 1965 and 1971 (Cohort 2). A very small proportion of Cohort 2 (less than 0.2%) appear to have been born before or after these dates. The Cohort 1 interviews took place in 1992 and 1993 and the age of these subjects at interview was 44 years on average (range = 27–90). The Cohort 2 interviews took place between 1996 and 2000 and the age of these subjects at interview was 30 years on average (range = 24–39). Postcode data for the addresses provided by the participants updated to the time of last contact were used to determine location. There is some uncertainty about the date of the last address update, which may have a small effect on the subjects' ages at the relevant time, but any errors introduced from this source will be small in relation to the differences in age between the three groups. Subjects' postcodes were categorized into urban, suburban or rural zones, as described in Table 1. Using these three categories as an ordinal variable, the polychoric correlations by zygosity and age group were estimated and models of genetic and nongenetic sources of variation fitted using Mx 1.57 (Neale et al., 2003).

Table 2

Breakdown of Numbers of Individual Subjects by Place-of-Residence Categories, for the Three Age-Based Groups of Study Participants

	Urban	Suburban	Nonurban
Born 1965–1971	809 (12%)	2589 (41%)	2998 (47%)
Born 1951–1964	476 (13%)	1485 (41%)	1682 (46%)
Born 1893–1951	350 (10%)	1475 (42%)	1702 (48%)

Results

Using the definitions of three residential zones as shown in Table 1, approximately 10% of the individuals with postcode data lived in the urban zone, 40% in the suburban zone, and 45% in the nonurban zone (Table 2). The polychoric correlations by twin-pair zygosity are shown in Table 3. These correlations were high and equal for MZ and DZ same-sex pairs in the youngest group (born 1965 to 1971), and the DZ correlations were lower (first for women, and then for men also) in the older groups. This suggested an initial predominance of shared environmental effects, with genetic variation becoming important later in life. The DZ opposite-sex correlations were lower than the same-sex pair correlations in the youngest group, suggesting that shared environmental effects are less important or different in opposite-sex pairs.

Table 3

Within-Pair Similarity of Place of Residence by Zygosity and Age

Zygosity	Polychoric correlations and number of pairs		
	Born 1965 to 1971	Born 1951 to 1964	Born 1950 or before
MZF	.63 (656)	.65 (512)	.57 (586)
MZM	.60 (484)	.56 (250)	.48 (266)
DZF	.64 (512)	.48 (317)	.40 (350)
DZM	.58 (427)	.63 (201)	.26 (128)
DZOS	.47 (658)	.49 (454)	.30 (352)

Note: The categories urban, suburban and nonurban were treated as an ordinal variable from which the correlations were estimated.

Table 4

Models of Genetic and Environmental Sources of Variation in Residential Location, Divided by Date of Birth

Models	–2LL	*df*	AIC	Δχ	Δ*df*	*p*	A	C	E
Youngest: Born 1965 to 1971									
ACE	**11936.485**	**6382**	**–827.515**				**12.7**	**49.2**	**38.1**
AE	12005.413	6383	–760.587	68.928	1	1.02E–16	68.8		31.2
CE	11939.68	6383	–826.32	3.195	1	.073864		58.2	41.8
Intermediate: Born 1951 to 1964									
ACE	**6780.177**	**3621**	**–461.823**				**22.0**	**40.3**	**37.7**
AE	6807.321	3622	–436.679	27.144	1	1.89E–07	67.7		32.3
CE	6785.987	3622	–458.013	5.81	1	.015935		56.3	43.7
Oldest: Born 1950 or before									
ACE	**6426.742**	**3505**	**–583.258**				**41.2**	**13.1**	**45.8**
AE	6428.815	3506	–583.185	2.073	1	.149926	55.9		44.1
CE	6441.321	3506	–570.679	14.579	1	.000134		44.5	55.5

These patterns of correlation led us to test models of sources of variation, with results shown in Table 4. These assume the same sources of variation for men and women; models which included sex limitation were also tested but showed no significant improvement in any of the age groups. The AE models were definitely rejected, and the CE models gave a marginally poorer fit than the ACE models in the youngest and the intermediate groups, while the CE model was rejected in the oldest group. Overall, the estimates of A (additive genetic), C (shared environment) and E (nonshared environment) under the ACE model showed an initial preponderance of C over A in the youngest group, shifting to a much greater A estimate than C in the oldest. This is illustrated in Figure 1.

Discussion

Our main finding is that effects of shared environment and of genetic factors on place of residence exist, but their relative importance varies with date of birth and therefore, presumably, age. In the youngest group, the urban/suburban/nonurban division was determined by shared and nonshared environmental factors. This can be seen from the correlations in the second column of Table 3, which are essentially identical for MZ and DZ same-sex pairs. It is interesting to note that in this group the estimated correlation was lower in opposite-sex pairs than for either male or female same-sex DZ pairs, suggesting that there are aspects of the environment which have been shared less by brother–sister pairs than by brothers with brothers or sisters with sisters. Shared environment may of course include peer interactions as well as the effects of parental characteristics on children, and these may well be sex-specific. However the model-fitting did not show significant sex limitation of either genetic or environmental factors, possibly because of lack of power using categorical variables.

Genetic factors become more important with increasing age. In the group born between 1951 and 1964, who would have been aged between 28 and 42 years at the relevant time, the genetic contribution appears, from the pattern of correlations, to be greater for women than for men. However, this was not significant in the comparison of models with and

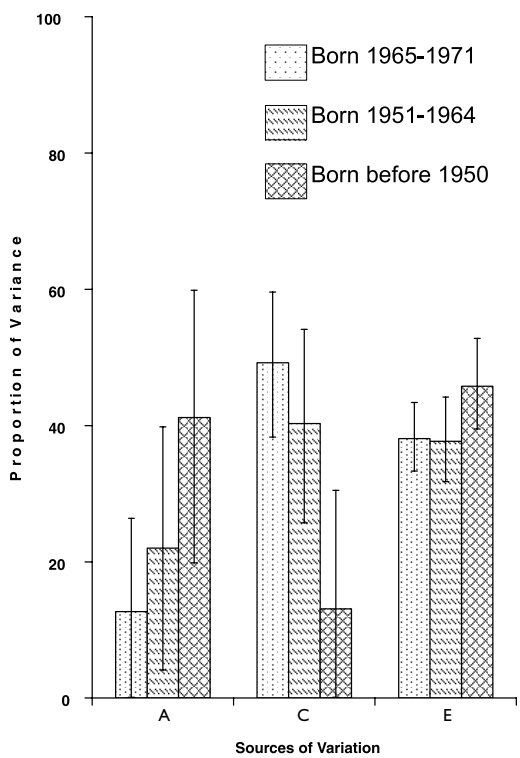

Figure 1

Proportions of variation in residential location attributable to additive genetic (A), shared environmental (C), and nonshared environmental (E) sources, in three groups divided by date of birth. Error bars show 95% confidence intervals of the estimates.

without sex limitation. In the oldest group, born before 1951 and aged between 41 and 90 years in 1992/1993, the genetic contribution to place of residence dominates in both men and women.

The proportion of variation due to shared environmental effects at the younger ages, and genes at the older ages, was surprisingly high. Over half the variation could be ascribed to these sources, even though the choice of residence would generally have involved both the twin and their spouse or other long-term partner (who would bring their own familial and genetic background to the decision). This may be due to assortative mating for social status, birthplace or parental residence, or other factors relevant to the urban/suburban/nonurban lifestyle choice. Such a pattern of assortative mating by place of residence at the time of partner choice would, in the absence of other factors promoting movement and exogamy, amplify population genetic variance for residential choice.

One explanation for the differences in sources of variation between the age groups is that as people grow older, they are more likely to make choices based on their own underlying preferences and natures and less likely to be affected by their family's background. A similar phenomenon, with an earlier age of change, has been proposed for the social attitude dimension of conservatism (Eaves et al., 1997). Because we only studied one time-point for each subject, it is not formally possible to distinguish age-related effects from secular trends, but the proportion of people living within each of the three zones was constant across the three age groups.

The criteria we used to define the three zones of urban, suburban and nonurban residence were based on subjective knowledge of the nature of Australia's cities rather than on, for example, census data on population density. Nevertheless, this definition gave some interesting insights. Any future work on this topic would benefit from a more rigorous approach to the definition of zones, or the use of population density as a continuous variable.

Factors affecting residential location and the choices which people make about this have a significance which extends beyond the real estate industry and government planning agencies. Urbanization and the factors which drive or retard it have a number of health, social and perhaps evolutionary implications. There is contemporary evidence that urbanization affects the prevalence of depression and psychosis (Sundquist et al., 2004), alcohol and drug dependence (Sundquist & Frank, 2004), and coronary heart disease (Sundquist et al., 2004). Although these findings may reflect social deprivation rather than population density, or perhaps a tendency of mentally disturbed individuals to move to inner-city areas, we cannot exclude an effect of urbanization and associated stresses on psychiatric and even physical morbidity. There is considerable information on the link between urbanization and psychosis, including evidence for a vulnerability period during childhood and early adolescence and dose-related risk (Pedersen & Mortensen, 2001).

In the past, when infectious disease and epidemics were a much greater cause of mortality, living in the densely populated cities probably increased mortality. Certainly outbreaks of plague prompted those who could do so to leave the cities. Offsetting this, from Dick Whittington's London to contemporary China, enterprising or desperate people have moved from the country to large cities because of their perception of better economic opportunities.

It is unlikely that such internal migration is random with respect to genes influencing important behavioral dimensions. Obviously the balance is affected by many complex and competing factors to which we now have to add a significant contribution from genetic differences between people. The mediating or modifying variables through which this is expressed are still unknown, and will repay further study. In particular, it would be interesting to test the generalizability of our findings to other countries that are demographically either similar to or different from Australia. The existence of twin registries in countries of comparatively recent European settlement such as Australia, Canada and the United States; many countries in Western Europe and Scandanavia; and Asian countries including China, Korea and Japan (Busjahn, 2002) will facilitate further testing and comparisons, and more detailed examination of our hypothesis.

Acknowledgments

The surveys from which the address data were obtained were supported by the National Institutes of Health (AA04535, AA07535), and the Australian Twin Registry is supported by the National Health and Medical Research Council. This analysis was prompted by a question from Professor Irving Gottesman, University of Minnesota.

References

Baker, L. A., Treloar, S. A., Reynolds, C. A., Heath, A. C., & Martin, N. G. (1996). Genetics of educational attainment in Australian twins: Sex differences and secular changes. *Behavior Genetics, 26*, 89–102.

Busjahn, A. (2002). Twin registers across the globe: What's out there in 2002? *Twin Research, 5,* v–vi.

Eaves, L. J., Martin, N. G., Heath, A. C., Schieken, R., Meyer, J., Silberg, J., Neale, M. C., & Corey, L. A. (1997). Age changes in the causes of individual differences in conservatism. *Behavior Genetics, 27,* 121–124.

Heath, A. C., Bucholz, K. K., Madden, P. A., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., Statham, D. J., Dunne, M. P., Whitfield, J. B., & Martin, N. G. (1997). Genetic and environmental contributions to alcohol dependence risk in a national twin sample: Consistency

of findings in women and men. *Psychological Medicine, 27,* 1381–1396.

Jardine, R., & Martin, N. G. (1984). Causes of variation in drinking habits in a large twin sample. *Acta Geneticae Medicae et Gemellologiae, 33*, 435–450.

Kendler, K. S., Martin, N. G., Heath, A. C., & Eaves, L. J. (1995). Self-report psychiatric symptoms in twins and their nontwin relatives: Are twins different? *American Journal of Medical Genetics (Neuropsychiatric Genetics), 60,* 588–591.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Nelson, E. C., Heath, A. C., Bucholz, K. K., Madden, P. A., Fu, Q., Knopik, V., Lynskey, M. T., Lynskey, M. T., Whitfield, J. B., Statham, D. J., & Martin, N. G. (2004). Genetic epidemiology of alcohol-induced blackouts. *Archives of General Psychiatry, 61*, 257–263.

Pedersen, C. B., & Mortensen, P. B. (2001). Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. *Archives of General Psychiatry*, *58*, 1039–1046.

Sundquist, K., & Frank, G. (2004). Urbanization and hospital admission rates for alcohol and drug abuse: A follow-up study of 4.5 million women and men in Sweden. *Addiction, 99,* 1298–1305.

Sundquist, K., Frank, G., & Sundquist, J. (2004). Urbanization and incidence of psychosis and depression: Follow-up study of 4.4 million women and men in Sweden. *British Journal of Psychiatry, 184*, 293–298.

Sundquist, K., Malmstrom, M., & Johansson, S. E. (2004). Neighbourhood deprivation and incidence of coronary heart disease: A multilevel study of 2.6 million women and men in Sweden. *Journal of Epidemiology and Community Health, 58,* 71–77.