# Progress in the Molecular-Genetic Study of Intelligence

**Danielle Posthuma and Eco J.C. de Geus**

*Department of Biological Psychology and Center for Neurogenomics and Cognitive Research, Vrije Universiteit, Amsterdam, The Netherlands*

**ABSTRACT—***The past decade has seen a major shift in the genetic study of human intelligence; where classic studies aimed to quantify the heritability of intelligence, current studies aim to dissect this heritability into its molecular-genetic components. Five whole-genome linkage scans have been published in the past year, converging on several chromosomal (or genomic) regions important to intelligence. A handful of candidate genes, some of which lie in these genomic regions, have shown significant association to intelligence and the associations have been replicated in independent samples. Finding genes brings us closer to an understanding of the neurophysiological basis of human cognition. Furthermore, when genes are no longer latent factors in our models but can actually be measured, it becomes feasible to identify those environmental factors that interact and correlate with genetic makeup. This will supplant the long nature–nurture debate with actual understanding.*

**KEYWORDS—***cognition; quantitative trait locus; gene–environment interaction; gene–environment correlation*

Individual performance on a single aspect of cognitive ability is highly predictive of performance on other aspects of cognitive ability. About 40% of the variance in multiple aspects of cognitive ability can be accounted for by a single general-intelligence factor. Genetic analyses of data obtained from monozygotic (from a single egg) and dizygotic (from different eggs) twin pairs indicate that this general-intelligence factor is highly heritable. Estimates range from 40% in childhood to 80% in late adulthood.[1] In keeping with the postulate of a general-intelligence factor, substantial genetic overlap is found not only between different aspects of cognitive ability (such as reading ability, working memory, and attention) but also between different biological levels of cognitive ability, such as brain size or protein levels. This observation has led to the recent formulation of the "generalist genes" hypothesis, which states that the same genes affect multiple cognitive abilities (Plomin and Kovas, 2005).

In recent years, molecular-genetic studies have indeed started to identify these genes, with the full awareness that they are "polygenes"—i.e., they may each explain only a very small part of the variation in one or more cognitive abilities. The generalist-genes hypothesis also implies that cognitive disabilities are the extremes of normally distributed dimensions of cognitive abilities. Therefore, exactly those genes that have been associated with normal cognitive abilities could provide important clues to underlying mechanisms of milder but more prevalent forms of impaired cognitive functioning, like reading disorder, dyslexia, and attention deficit hyperactivity disorder, or even the severe cognitive deficits seen in autism and schizophrenia.

## GENE-FINDING STRATEGIES

To identify genes underlying genetic variation in intelligence, two main strategies are available: linkage analysis and candidate-gene association studies. In linkage analysis, a number of DNA markers of known location, evenly dispersed throughout the entire set of chromosomes (the genome), are measured in genetically related individuals. DNA markers can be mutations in a single base pair (the smallest unit of the DNA helix; such mutations are called single nucleotide polymorphisms, SNPs) or a variable number of repeats of two or more base pairs (microsatellites). They need not be part of a functional gene—they are merely landmarks in the genome. For each DNA marker, evidence for linkage to a particular trait, like cognitive ability, is obtained through statistical procedures that trace how often the trait and the DNA marker are jointly passed along in familial lineages (cosegregation). If such a cosegregation of a DNA

Address correspondence to Danielle Posthuma, Vrije Universiteit Amsterdam, Department of Biological Psychology, van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands; e-mail: danielle@psy.vu.nl.

[1]The heritability of a trait is a ratio of the genetic variance to the total variance of that trait, so changes in heritability can be the result of changes in environmental variation or in genetic variation. The latter in turn may be caused by the same genes having differential effects across different ages (gene amplification) or by genes turning on and off at certain points in development (gene emergence).

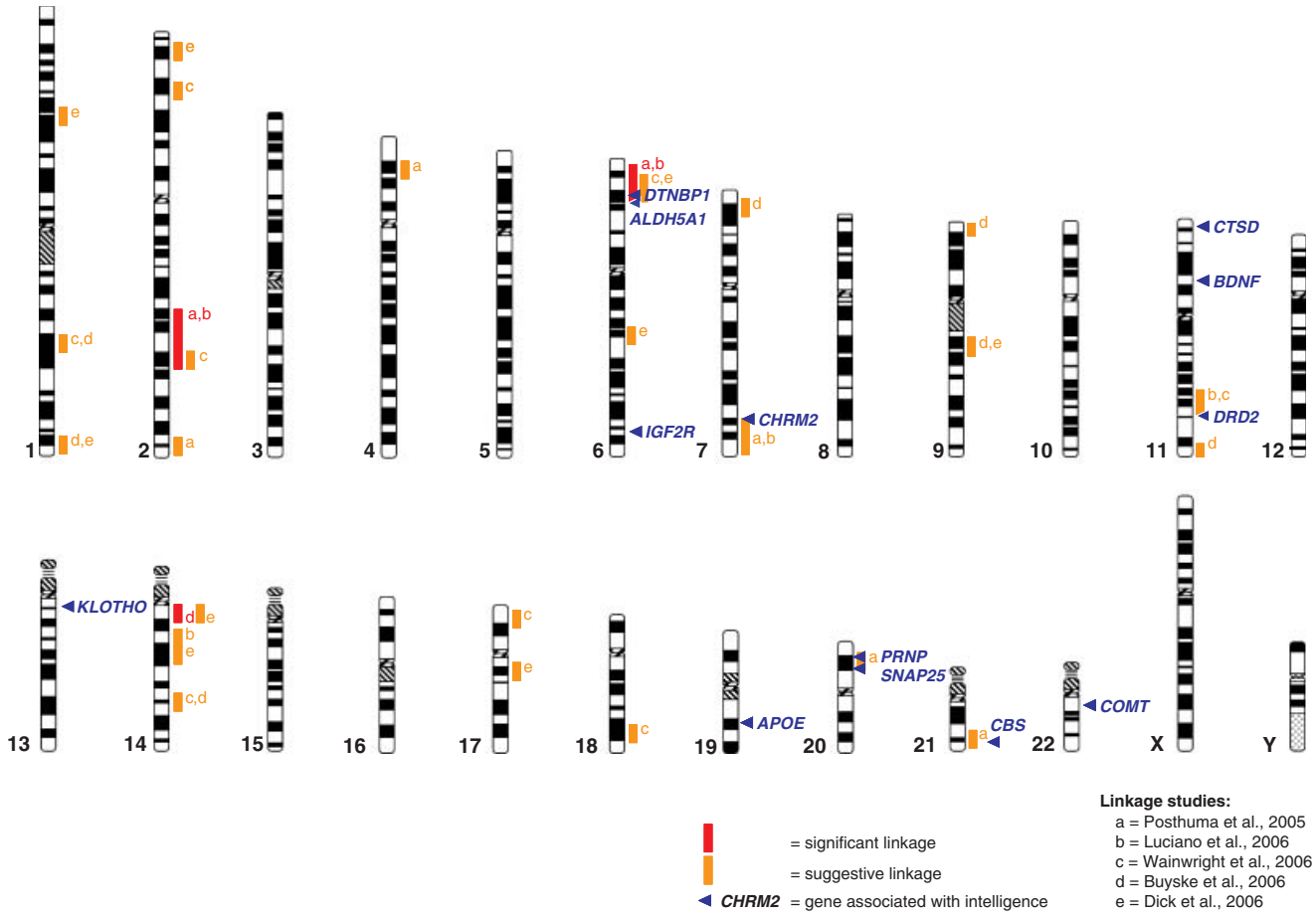


**Fig. 1.** Ideogram (chromosome map) of the human genome, showing which regions of chromosomes are likely to contain genes for intelligence (marked in red), as based on the five linkage studies for intelligence that have been conducted to date. It also shows the chromosomal regions of all the genes that have been associated with intelligence so far (marked in orange).

marker and a trait can be established with sufficient statistical confidence, then one or more genes in the vicinity of the marker are possibly involved in trait similarity among individuals, because genetic material on chromosomes is passed on in chunks. Linkage analysis thus serves to detect the regions (called quantitative trait loci) of the genome where genetic variants with a quantitative effect on the trait must be located.

The first whole-genome linkage scan for intelligence was published in 2005 (Posthuma et al., 2005), and four more studies have been published since (Buyske et al., 2006; Dick et al., 2006; Luciano et al., 2006; Wainwright et al., 2006). The sample used in the first study consisted of a Dutch sample (159 sibling pairs) and an Australian sample (475 sibling pairs). Results indicated two areas of significant linkage to general intelligence—one on the long arm (denoted as q) of chromosome 2 (i.e., 2q) and one on the short arm (denoted as p) of chromosome 6 (i.e., 6p)—and several areas of suggestive linkage (an additional site on 2q, as well as areas on 4p, 7q, 20p, and 21p). The chromosome-2 area has been implicated in linkage scans for autism and dyslexia, while the chromosome-6 area is the main linkage area for reading ability and dyslexia.

Two studies with a partly overlapping sample confirmed the importance of the areas on chromosomes 2 and 6 for specific aspects of intelligence (Luciano et al., 2006) as well as for academic achievement, which is different from IQ score but predicts IQ very well (correlation around 0.6; Wainwright et al., 2006). The Luciano et al., (2006) study additionally showed that both word recognition and IQ were linked to chromosome 2, confirming the notion of the same genes influencing different aspects of cognitive ability (Plomin and Kovas, 2005). A completely independent study by Dick et al. (2006) using data collected as part of the Collaborative Study on the Genetics of Alcoholism (COGA) also confirmed linkage of intelligence to the chromosome-6 area. A second scan based on that dataset (Buyske et al., 2006) found strong evidence for linkage of specific cognitive abilities to chromosome 14, an area that showed suggestive evidence for linkage in three of the five linkage studies (see Fig. 1). Although the COGA dataset has been selected for alcohol dependence and may thus not be representative of the general population, Dick et al. (2006) showed that alcohol dependence explained less than 1% of the variance in IQ scores. Moreover, a correction for the selection strategy used

(i.e., alcohol dependency) did not change the results significantly.

In association analysis, candidate genes are selected based on existing neuroscientific evidence. Variation in these genes (i.e. alleles) is measured and tested for association with intelligence. These measured allelic variants can either be functional variants that change the gene's effect on intelligence or variants that do not alter the gene effect but are in close vicinity to the true (but unmeasured) functional variant. Case-control association studies—in which, for example, a sample of subjects with high IQ scores (cases) is compared with a sample of subjects with lower IQ (controls)—have the highest statistical power but may provide spurious associations as a result of the use of stratified samples. If many alleles show a different distribution across cases and controls for reasons unrelated to IQ (e.g., due to unbalanced ethnicity in the cases and controls or due to nonrandom or assortative mating in which mate selection is based on resemblance for IQ), all of these alleles will show a spurious statistical association to IQ. Family-based association studies are statistically less powerful than case-control studies but control for these effects of population stratification, as allelic association is tested exclusively within members of the same family.

In association analysis, the choice of candidate genes is crucial. It is usually based on prior knowledge of the gene's involvement in biological functions relevant to intelligence, such as neurophysiological systems known to influence human memory and cognition. Candidate genes can also be selected based on results from animal studies in which the genes have been shown to influence performance on tests of learning and memory. Finally, a large number of genes have been associated with mental retardation (Inlow and Restifo, 2004). Although most mental retardation is monogenetic—i.e., caused by a severe mutation in a single gene only—subtler variation in these genes might influence normal variation in intelligence.

In 1998, Chorney et al. published the first genetic association for normal variation in intelligence, effectively marking the advent of the current era of molecular genetics. They tested 37 markers on the long arm of chromosome 6 for association with IQ in young children. Chromosome 6 was chosen basically because it was the first large chromosome expected to be sequenced by the Human Genome Project, but also because it contained the insulin-like growth factor-2 receptor (IGF2R) gene, which was previously found to be active in brain regions involved in learning and memory. The IGF2R gene was indeed shown to be associated with high IQ in children (Chorney et al., 1998). Four years later, however, this gene was not found to be associated with IQ in an enlarged sample (Hill, Chorney, Lubinski, Thompson, & Plomin, 2002). The same group of researchers continued with a major effort to identify further genetic variants using a whole-genome association approach. This approach resembles linkage more than it resembles the candidate-gene approach in that it aims to identify quantitative trait loci throughout the entire genome, although based on association of the trait to DNA markers rather than cosegregation of the trait and the markers within families. The only association with intelligence in young children that survived rigorous adjustment for multiple testing was a significant association of a functional variant in the aldehyde dehydrogenase 5 family, member A1 (ALDH5A1) gene that causes increased activity of the aldehyde dehydrogenase enzyme (Plomin et al., 2004).

Because heritability increases from 40% in childhood to 80% in adulthood, young children may not be the optimal population for gene finding. At the same time, samples including the elderly may inadvertently confound genetic effects on intelligence with genetic effects on the speed of cognitive aging. The ε4 variant of the apolipoprotein-E (APOE) gene, for instance, has been associated with cognitive deterioration in old age but does not seem to be unambiguously related to cognitive ability or intelligence in children or adults (Deary et al., 2002). Its effect on cognitive ability in the elderly may well derive entirely from the role of this gene in the development of Alzheimer's dementia.

As summarised in Figure 1, various other genes have shown significant association to intelligence.[2] These include the CTSD gene, the PRNP gene, the DRD2 gene, the CBS gene, the BDNF gene, and the COMT gene. These genes have been associated not only with intelligence but with a wide range of cognitive abilities, mostly related to attention (DRD2) and working memory (COMT, BDNF), or to pathological brain states like Alzheimer's dementia (CTSD, BDNF, APOE, PRNP, CBS, COMT), aphasia (APOE, PRNP), and mental retardation (APOE, CBS). To understand the underlying mechanisms linking individual differences in specific cognitive abilities to intelligence or to brain pathology we should adopt a multivariate approach in which we test, for example, whether the same set of genes influences both the specific cognitive ability and an index of brain pathology. This multivariate approach has the added advantage of increasing the power of gene-finding studies.

It is important to note that most of the associations in Figure 1 have proven difficult to replicate. Poor replication of an initially promising association or linkage result is a common concern in the molecular-genetic study of complex brain functioning. Ideally, associations with candidate genes are replicated in independent samples and these candidate genes should also surface in whole-genome searches. To date, only one such genetic association is known for intelligence. In 2003, Comings et al. reported that a variant of the cholinergic muscarinic receptor 2 (CHRM2) gene explained 1% of the variance in full-scale IQ. Two years later, suggestive linkage for intelligence was found on the long arm of chromosome 7 (7q), right above the CHRM2 gene (Posthuma et al., 2005). Subsequently, Gosso et al. (2006) replicated the association between the CHRM2 gene and intelligence in a combined sample of Dutch 12-year-olds and Dutch

[2] Although many X-linked genes have been related to mental retardation, most of these genes are monogenetic genes as detailed earlier. These genes are extensively reviewed in Inlow and Restifo, 2004, but are not included in Figure 1.

adults. Here the gene explained 2% of the total variance in full-scale IQ. Although both Comings et al. (2003) and Gosso et al. (2006) did not include functional variants of the CHRM2 gene, the variants they tested were in the same region of this gene, suggesting that functional variants within that region are important to intelligence.

In summary, the last decade has started to see the now well-established heritability of cognitive ability being dissected into its molecular-genetic elements. Association studies have yielded a handful of polygenes of small effect, of which only one has shown replicated association with intelligence. Recent linkage studies for intelligence will refocus association studies to those genes that lie in areas of genetic linkage, a strategy that has proven successful for the CHRM2 gene. These polygenes bring us one step closer to understanding the biological basis of intelligence, although we still face the daunting task of charting the exact route from genetic variation to variation in brain function and on to individual differences in intelligence.

## GENE–ENVIRONMENT INTERACTION AND CORRELATION

Complex traits such as intelligence are expected to be influenced at least partly by interactions between genes and environmental factors, and genetic variation may not be evenly distributed across all environments. The gene-finding strategies reviewed earlier completely ignore gene–environment interaction and gene–environment correlation. This may prove a costly oversimplification, and a main improvement in future gene-finding strategies would be to bring the environment into the equation.

Crucial to the estimation of the heritability of intelligence is the observation that monozygotic twins, who are effectively genetically identical, correlate about .60 to .80 on tests of intelligence, whereas dizygotic twins, who share only half of their genetic material, correlate around half of that. Under the assumption that such a pattern of twin correlations is explained by a simple addition of the separate effects of genes and shared and nonshared environmental factors, this yields a heritability of intelligence in adulthood of 60 to 80%. However, the same pattern of twin correlations may also partly result from interactions between genetic and shared environmental effects, or by correlations between genetic variation and environmental factors.

The interaction between genes and shared environment will mimic genetic effects in statistical models. It comes about when a favorable—e.g., intellectually stimulating—family environment has more impact on individuals with a particular genetic make-up than it does on others ("fertile ground"). Environmental mediation of genetic effects for intelligence has for example been shown for socioeconomic status (SES) and parental education, in the sense that the heritability of intelligence is lower for subjects with low SES or when their parents have received less years of education, and higher for persons with high SES or when parents received many years of education.

Gene–environment correlation may occur when parents transmit not only their genes but also their environment, a mechanism known as cultural transmission. For example, parents who are at the high end of the intelligence scale may transmit both the genetic variants associated with higher intelligence and a home environment that provides easy access to intellectual knowledge (such as the availability of books, intellectual discussions at the dinner table, and focus on obtaining high grades at school). Such gene–environment correlation tends to increase the DZ correlation, while the MZ correlation remains the same. Another form of gene–environment correlation occurs when a genetic makeup that favors high intelligence will also more often result in the selection of a societal niche that is conducive to the development of intelligence (such as smarter kids being admitted to the more advanced forms of schooling). Even more complex models allow reciprocal causation between intelligence and environmental factors, resulting in strong correlations between genetic endowment and favorable environmental conditions (Dickens and Flynn, 2001).

A final well-known form of actively induced gene–environment correlation is assortative mating, which not only affects the presence of environmental factors in a given person but also affects resemblance in traits among that person and his or her offspring. Assortative mating is reflected in a spousal correlation greater than zero, and is known to indeed exist for intelligence, where spousal correlations range around 0.30. When smart mothers more often elect smart fathers as mates (and vice versa), this will increase the resemblance between parents and offspring as well as that between siblings and that between dizygotic twins. In twin studies, this may conceal the presence of non-additive genetic effects (gene–gene interactions or genetic dominance) and overestimate the influence of additive genetic factors.

If we assume that the high heritability of intelligence is at least partly explained by complex mechanisms such as gene–gene interaction, gene–environment interaction, and gene–environment correlation, gene finding without taking these complex mechanisms into account will prove very difficult (as is already proving to be the case). Fortunately, as genes are no longer "latent factors" in our models but can actually be measured, investigating genetic effects while allowing for the interplay between genes and environmental factors has now become a realistic goal. Measured candidate genes, such as the CHRM2 gene, allow testing the effects of genes under different "experimental" environmental conditions, using simple designs such as comparing the effect of variation in the gene in groups of children with high- or low-educated parents.

By identifying, replicating, and functionally describing more polygenes like CHRM2, the next era of molecular-genetic research on intelligence will be better enabled to consider the complex interplay of genes and environmental influences. Ultimately, we may be able to supplant the long nature–nurture

debate on intelligence with actual understanding of the biological processes in brain development that lead to individual differences in cognitive abilities and disabilities alike. Such understanding will have clear practical benefits for education and learning theories. Genes influencing psychometric IQ are just tiny pieces in this very complex puzzle. However, they are edge pieces. As with any complex puzzle, it may be a good idea to start laying it from the edges.

**Recommended Reading**

Dickens, W.T., & Flynn, J.R. (2001). (See References)

Goldberg, T.E., & Weinberger, D.R. (2004). Genes and the parsing of cognitive processes. *Trends in Cognitive Sciences*, *8*, 325–335.

Gray, J.R., & Thompson, P.M. (2004). Neurobiology of intelligence: Science and ethics. *Nature Reviews Neuroscience*, *5*, 471–482.

Kovas, Y., & Plomin, R. (2006). Generalist genes: Implications for the cognitive sciences. *Trends in Cognitive Sciences*, *10*, 198–203.

Plomin, R., & DeFries, J.C. (1998). The genetics of cognitive abilities and disabilities. *Scientific American*, *278*(5), 62–69.

***Acknowledgments***—Supported by the Human Frontiers of Science Program (grant number RG0154/1998-B), and the Netherlands Organization for Scientific Research (NWO/MaGW Vernieuwingsimpuls 016-065-318).

## REFERENCES

Buyske, S., Bates, M.E., Gharani, N., Matise, T.C., Tischfield, J.A., & Manowitz, P. (2006). Cognitive traits link to human chromosomal regions. *Behavioral Genetics*, *36*, 65–76.

Chorney, M.J., Chorney, K., Seese, N., Owen, M.J., McGuffin, P., Daniels, J., Thompson, L.A., Detterman, D.K., Benbow, C.P., Lubinski, D., Eley, T.C., & Plomin, R. (1998). A quantitative trait locus associated with cognitive ability in children. *Psychological Science*, *9*, 159–166.

Comings, D.E., Wu, S., Rostamkhani, M., McGue, M., Lacono, W.G., Cheng, L.S., & MacMurray, J.P. (2003). Role of the cholinergic muscarinic 2 receptor (*CHRM2*) gene in cognition. *Molecular Psychiatry*, *8*, 10–11.

Deary, I.J., Whiteman, M.C., Pattie, A., Starr, J.M., Hayward, C., Wright, A.F., Carothers, A., & Whalley, L.J. (2002). Cognitive change and the APOE epsilon 4 allele. *Nature*, *418*, 932.

Dick, D.M., Aliev, F., Bierut, L., Goate, A., Rice, J., Hinrichs, A., Bertelsen, S., Wang, J.C., Dunn, G., Kuperman, S., Schuckit, M., Nurnberger, J., Jr., Porjesz, B., Beglieter, H., Kramer, J., & Hesselbrock, V. (2006). Linkage Analyses of IQ in the Collaborative Study on the Genetics of Alcoholism (COGA) Sample. *Behavioral Genetics*, *36*, 77–86.

Dickens, W.T., & Flynn, J.R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, *108*, 346–369.

Gosso, M.F., van Belzen, M., de Geus, E.J.C., Polderman, J.C., Heutink, P., Boomsma, D.I., & Posthuma, D. (2006). Association between the CHRM2 gene and intelligence in a sample of 304 Dutch families. *Genes, Brain and Behavior*. Retrieved July 3, 2006, from http://www.blackwell-synergy.com/doi/pdf/10.1111/j.1601-183x.2006.00211.x

Hill, L., Chorney, M.J., Lubinski, D., Thompson, L.A., & Plomin, R. (2002). A quantitative trait locus not associated with cognitive ability in children: A failure to replicate. *Psychological Science*, *13*, 561–562.

Inlow, J.K., & Restifo, L.L. (2004). Molecular and comparative genetics of mental retardation. *Genetics*, *166*, 835–881.

Luciano, M., Wright, M.J., Duffy, D.L., Wainwright, M.A., Zhu, G., Evans, D.M., Geffen, G.M., Montgomery, G.W., & Martin, N.G. (2006). Genome-wide Scan of IQ Finds Significant Linkage to a Quantitative Trait Locus on 2q. *Behavioral Genetics*, *36*, 45–55.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, *131*, 592–617.

Plomin, R., Turic, D.M., Hill, L., Turic, D.E., Stephens, M., Williams, J., Owen, M.J., & O'Donovan, M.C. (2004). A functional polymorphism in the succinate-semialdehyde dehydrogenase (aldehyde dehydrogenase 5 family member A1) gene is associated with cognitive ability. *Molecular Psychiatry*, *9*, 582–586.

Posthuma, D., Luciano, M., de Geus, E.J.C., Wright, M.J., Slagboom, P.E., Montgomery, G.W., Boomsma, D.I., & Martin, N.G. (2005). A genome-wide scan for intelligence identifies quantitative trait loci on 2q and 6p. *American Journal of Human Genetics*, *77*, 318–326.

Wainwright, M.A., Wright, M.J., Luciano, M., Montgomery, G.W., Geffen, G.M., & Martin, N.G. (2006). A linkage study of academic skills defined by the Queensland Core Skills Test. *Behavioral Genetics*, *36*, 56–64.