Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism

Avshalom Caspi*†‡, Benjamin Williams*, Julia Kim-Cohen§, Ian W. Craig*, Barry J. Milne*, Richie Poulton¶, Leonard C. Schalkwyk*, Alan Taylor*, Helen Werts*, and Terrie E. Moffitt*†

*Medical Research Council Social, Genetic, and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London SE5 8AF, England; †Departments of Psychology and Neuroscience, and Psychiatry and Behavioral Sciences, and Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708-0086; §Department of Psychology, Yale University, New Haven, CT 06520; and ¶Dunedin School of Medicine, University of Otago, P.O. Box 913, Dunedin, New Zealand

Edited by William T. Greenough, University of Illinois at Urbana–Champaign, Urbana, IL, and approved October 9, 2007 (received for review May 14, 2007)

Children's intellectual development is influenced by both genetic inheritance and environmental experiences. Breastfeeding is one of the earliest such postnatal experiences. Breastfed children attain higher IQ scores than children not fed breast milk, presumably because of the fatty acids uniquely available in breast milk. Here we show that the association between breastfeeding and IQ is moderated by a genetic variant in *FADS2*, a gene involved in the genetic control of fatty acid pathways. We confirmed this gene–environment interaction in two birth cohorts, and we ruled out alternative explanations of the finding involving gene–exposure correlation, intrauterine growth, social class, and maternal cognitive ability, as well as maternal genotype effects on breastfeeding and breast milk. The finding shows that environmental exposures can be used to uncover novel candidate genes in complex phenotypes. It also shows that genes may work via the environment to shape the IQ, helping to close the nature versus nurture debate.

cognitive development | gene environment interaction

For 100 years, the IQ has been at the heart of scientific and public debates about nature versus nurture (1–3). Twin studies document that differences between individuals' IQs are under strong genetic influence, but twin studies also attest to the existence of nongenetic, environmental influences on IQ, particularly for young children (4, 5). In the past 5 years, the nature versus nurture debate has shifted toward interest in how both nature and nurture work together (6). An integral part of this new focus is research that tests how genetic differences moderate the effects of environmental influences on individuals' health and behavior (7). Here we report replicated evidence that a measured genotype can moderate response to an environmental influence on children's IQ. We began our investigation of gene–environment interaction in IQ by selecting for study an environmental factor thought to influence neurodevelopment and known to predict IQ. We selected being fed breast milk (hereafter breastfeeding) as the environmental exposure because the biological processes underlying its benefits for the developing brain are increasingly well understood (8). A gene involved in these putative biological processes would be a good candidate for framing a gene–environment interaction hypothesis (9). Thus, selecting breastfeeding as the environmental exposure allowed us to nominate a novel candidate gene for this study of IQ.

Breastfeeding is thought to influence brain development through nutritional processes involving fatty acids (10). The predominant long-chain polyunsaturated fatty acids (LC-PUFAs) present in human milk, but not in cow's milk or most infant formulas, are docosahexaenoic acid (DHA; 22:6n-3) and arachidonic acid (AA or ARA; 20:4n-6) (11). Substantial amounts of DHA and AA accumulate in the human brain during the first postnatal months (12), and infants who are breastfed have higher concentrations of DHA and AA than infants fed unsupplemented formulas (13, 14). Evidence, in general, is consistent with the hypothesis that LC-PUFAs in breast milk may enhance cognitive development (15). In humans, children who are breastfed have higher IQs than children not fed breast milk (16, 17), and this advantage persists into adulthood (17). Although breastfeeding in contemporary, industrialized nations is associated with higher social class, IQ differences between breastfed children and children not fed breast milk remain significant in most observational studies even after adjustments for class-related confounding factors (16, 17). However, the essentiality of fatty acids cannot be inferred from such studies. Experimental studies, where more control can be achieved, show that animals that are fed diets deficient in n-3 fatty acids exhibit neuronal deficits, including memory, sensory, and visual abnormalities (18). DHA supplementation in rodents and nonhuman primates leads to increased brain DHA concentrations and enhanced performance on a wide variety of learning, memory, and problem-solving tasks (19–21). LC-PUFAs are thought to be important for cognitive development because they are required for efficient neurotransmission (22) and are involved in neurite outgrowth, dendritic arborization, and neuron regeneration after cell injury (23). This putative biological pathway led us to search the KEGG database (24) for genes involved in LC-PUFA metabolism, reasoning that they might moderate the effect of breastfeeding on children's IQ.

This search led us to *FADS2*, an attractive candidate gene because of its role in the modification of dietary fatty acids. *FADS2*, located on chromosome 11q12.2, encodes the delta-6 desaturase that is the rate-limiting step on the metabolic pathway leading to AA and DHA production. *FADS2* gene expression is also regulated through end-product inhibition and dietary LC-PUFAs such as those available in breast milk (25). To the best of our knowledge, *FADS2* polymorphisms have not been studied in relation to breast milk or to IQ. As such, we selected two SNPs (rs174575 and rs1535) as candidate biomarkers because they provided the best combination of two desirable factors: (*i*) they had known linkage disequilibrium (LD) throughout the region, and (*ii*) they had minor allele frequencies sufficiently prevalent to permit their use in tests of gene–environment interaction. Specifically, using data from CEPH (Utah residents with ancestry from northern and western Europe) HapMap trios (26), we found that these SNPs showed strong LD throughout the promoter and intragenic region of *FADS2*. In addition, these SNPs exhibited strong LD into the promoter and intragenic region of

Author contributions: A.C., B.W., and T.E.M. designed research; A.C., B.W., J.K.-C., I.W.C., R.P., L.C.S., H.W., and T.E.M. performed research; A.C., B.J.M., L.C.S., and A.T. analyzed data; and A.C., B.W., J.K.-C., and T.E.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

‡To whom correspondence should be addressed. E-mail: a.caspi@iop.kcl.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0704292104/DC1.

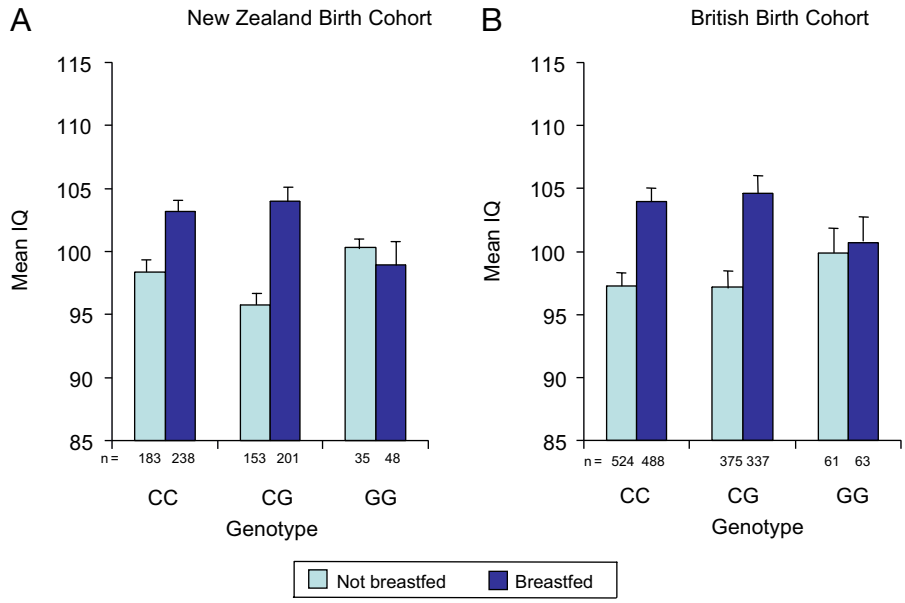


Fig. 1. The association between breastfeeding and IQ is moderated by a genetic polymorphism (rs174575) in the *FADS2* gene. In each cohort, we estimated a hierarchical regression model (ordinary least squares) with main effects for genotype (C carriers vs. GG homozygotes) and environment (not breastfed vs. breastfed) followed by a multiplicative genotype × environment interaction term, with covariate adjustment for socioeconomic status. In the Dunedin cohort (*A*), the effect of breastfeeding was significant ($t = 4.67$, $P < 0.001$), the effect of rs174575 was not significant ($t = 0.32$, $P = 0.75$), and the interaction term was significant ($t = 2.11$, $P = 0.035$). In the E-risk cohort (*B*), the effect of breastfeeding was significant ($t = 3.20$, $P < 0.001$), the effect of rs174575 was not significant ($t = 1.82$, $P = 0.42$), and the interaction term was significant ($t = 2.37$, $P = 0.018$).

FADS1, a highly similar gene that borders the 5′ region of *FADS2* and that is also involved in fatty acid metabolism, encoding the delta-5 desaturase (27). By genotyping these two tag SNPs (identified with the Gabriel method in Haploview v3.2 on release 20 of CEPH HapMap data), we could obtain maximum information about a candidate locus that potentially moderates breastfeeding effects on IQ [see supporting information (SI) *Text* and SI Figs. 3–5 for details about candidate gene and marker selection strategy]. We then tested the hypothesis that the cognitive advantage associated with breastfeeding in humans is related to genetic differences in LC-PUFA metabolism, and we replicated this test in two birth cohorts.

Results

Consistent with previous reports, the difference in IQ test scores between breastfed children and those not breastfed was 5.6 and 6.3 IQ points in the Dunedin and E-risk cohorts, respectively. Genotype was not related to IQ in either cohort. (IQ means associated with the three rs174575 genotype groups, CC, CG, and GG, were 101.1, 100.4, and 99.5 in Dunedin and 100.5, 100.7, and 100.3 in E-risk; IQ means associated with the three rs1535 groups, AA, AG, and GG, were 101.2, 100.3, and 100.9 in Dunedin and 101.0, 100.4, and 99.3 in E-risk.)

Analyses revealed that rs174575 interacted with breastfeeding to influence IQ in both the Dunedin ($P = 0.035$) and E-risk ($P = 0.018$) cohorts (Fig. 1). There was a dominant effect of the C allele in response to breastfeeding. In Dunedin, breastfed children carrying the C allele showed a 6.4-IQ-point advantage relative to children not fed breast milk ($t = 6.35$, $P < 0.001$). In contrast, GG homozygotes neither gained an advantage from breastfeeding nor suffered a disadvantage from not being fed breast milk ($t = 0.50$, $P = 0.62$) (Fig. 1*A*). Turning to the E-risk cohort, we found that breastfed children carrying the C allele showed a 7.0-IQ-point advantage relative to children not fed breast milk ($t = 7.91$, $P < 0.001$), whereas GG homozygotes neither gained an advantage from breastfeeding nor suffered a disadvantage from not being fed breast milk ($t = 0.22$, $P = 0.83$) (Fig. 1*B*).

Four points are relevant for interpreting this replicated gene–environment interaction between rs174575 and breastfeeding in predicting IQ. First, it is important to rule out confounding by social class, because socioeconomic advantage is related to children's higher IQ, and in modern countries, socioeconomically advantaged women are more likely to breastfeed (Table 1). To rule out this potential confound, all significance tests reported here for the rs174575-breastfeeding interaction were conducted with covariate adjustment for social class (see SI Table 2). Second, it is important to rule out confounding by maternal IQ (28). We added statistical controls for measures of maternal cognitive ability (Table 1); the rs174575-breastfeeding interaction remained significant in both Dunedin ($P = 0.03$) and E-risk ($P = 0.03$) (see SI Table 2). Third, to interpret the interaction, it is necessary to rule out potential genotype effects on exposure to breastfeeding. Child genotype was not related to breastfeeding in either cohort; prevalence rates of breastfeeding associated with the three rs174575 genotype groups (CC, CG, and GG) were 56%, 57%, and 58% in Dunedin [χ^2 (2) = 0.10, $P = 0.95$] and 48%, 47%, and 51% in E-risk [χ^2 (2) = 0.30, $P = 0.86$].

Fourth, it is important to rule out potential genotype differences in intrauterine growth. Because small gestational age and lower birth weight have been linked to lower IQ (29, 30), a spurious rs174575-breastfeeding interaction could be produced if GG homozygotes differed from C-carriers in their intrauterine growth. However, there were no significant gestational age or birth weight differences between the genotype groups, in either cohort (Table 1).

We repeated all analyses using rs1535. We observed a significant interaction in the E-risk cohort ($P = 0.01$). Breastfed A-carriers had higher IQs than nonbreastfed A-carriers, whereas this advantage was not as pronounced among GG homozygotes (breastfed children with AA, AG, and GG genotypes had IQs of 104.6, 104.6, and 100.0, whereas nonbreasfted children had IQs of 97.7, 96.8, and 98.6). This interaction did not replicate in the Dunedin cohort ($P = 0.55$); breastfed children with AA, AG, and

Table 1. Comparison of children in two birth cohorts, grouped according to genotype (rs174575) and breastfeeding, on tested intelligence (IQ) and covariates

	rs174575 CC homozygotes		rs174575 CG heterozygotes		rs174575 GG homozygotes	
Samples and measures	Not breastfed	Breastfed	Not breastfed	Breastfed	Not breastfed	Breastfed
New Zealand (Dunedin) birth cohort	$n = 183$	$n = 238$	$n = 153$	$n = 201$	$n = 35$	$n = 48$
Children's IQ	98.4 (15.2)	103.2 (13.9)	95.8 (12.4)	104.0 (13.4)	100.3 (11.2)	98.9 (13.8)
Socioeconomic status (1 = low; 3 = high)*	1.9 (0.60)	2.0 (0.65)	1.8 (0.55)	2.1 (0.60)	1.9 (0.53)	1.9 (0.58)
Maternal cognitive ability†	97.1 (15.2)	102.4 (15.2)	96.0 (12.9)	103.6 (14.2)	100.2 (13.4)	103.3 (14.9)
Gestational age,‡ weeks	40.0 (1.7)	40.0 (1.5)	39.7 (2.0)	40.2 (1.5)	39.9 (1.7)	40.2 (1.6)
Birthweight,§ g	3,399 (535)	3,374 (491)	3,289 (609)	3,467 (462)	3,431 (450)	3,344 (347)
British (E-risk study) birth cohort	$n = 524$	$n = 488$	$n = 375$	$n = 337$	$n = 61$	$n = 63$
Children's IQ	97.3 (14.1)	104.0 (15.0)	97.2 (13.9)	104.6 (15.3)	99.9 (15.3)	100.7 (17.3)
Socioeconomic status (1 = low; 3 = high)*	1.7 (0.75)	2.3 (0.79)	1.7 (0.76)	2.3 (0.75)	1.9 (0.77)	2.4 (0.68)
Maternal cognitive ability†	95.1 (14.9)	105.0 (12.9)	98.5 (12.7)	105.2 (13.8)	91.8 (14.5)	102.6 (14.4)
Gestational age,‡ weeks	36.4 (2.6)	36.1 (2.9)	36.4 (2.2)	36.2 (3.1)	36.5 (2.8)	36.0 (3.6)
Birthweight,§ g	2,452 (517)	2,404 (572)	2,442 (485)	2,466 (550)	2,490 (462)	2,483 (690)

Entries in the table are means and standard deviations. IQ scores were standardized to $M = 100$ and SD = 15 in each cohort.

*In both the Dunedin and E-risk cohorts, genotype was not associated with social class ($P = 0.34$ and 0.23), breastfeeding was significantly associated with social class ($P < 0.001$), and there was no difference in the association between breastfeeding and social class by genotype ($P = 0.93$ and 0.77).

†In the Dunedin cohort, maternal cognitive ability was assessed with the SRA verbal test (56). In the E-risk cohort, mothers were administered the Wide Range Achievement Test (57). Scores were standardized to $M = 100$ and SD = 15. In both the Dunedin and E-risk cohorts, genotype was not associated with maternal IQ ($P = 0.39$ and 0.84), breastfeeding was significantly associated with maternal IQ ($P < 0.001$), and there was no difference in the association between breastfeeding and maternal IQ by genotype ($P = 0.85$ and 0.34).

‡In both the Dunedin and E-risk cohorts, genotype was not associated with gestational age ($P = 0.81$ and 0.78), breastfeeding was associated with gestational age in Dunedin ($P = 0.04$) but not in E-risk ($P = 0.11$), and there was no difference in the association between breastfeeding and gestational age by genotype ($P = 0.13$ and 0.99).

§In both the Dunedin and E-risk cohorts, genotype was not associated with birth weight ($P = 0.99$ and 0.32), breastfeeding was not associated with birth weight ($P = 0.13$ and 0.61), and there was no difference in the association between breastfeeding and birth weight by genotype ($P = 0.27$ and 0.39).

GG genotypes had IQs of 102.7, 103.6, and 102.7, respectively, whereas nonbreastfed children had IQs of 99.1, 96.0, and 98.4.

The interaction between children's rs174575 genotype and breastfeeding, replicated across cohorts, suggests that C-carrying children benefit from breast milk more than do GG homozygotes. An alternative hypothesis is that there are variations in human-milk composition that are related to maternal rs174575 genotype. If our findings reflected such a maternal genotype effect, we should find that, among breastfeeding mothers, rs174575 C-carrying mothers have children with higher IQs than GG mothers. To test this possibility, we collected DNA from the mothers of the E-risk children. (We were not able to collect DNA from the mothers of the Dunedin cohort.) There were no significant IQ differences among children fed breast milk as a function of maternal genotype (breast-milk-fed children of CC, CG, and GG mothers had IQs of 104.6, 103.4, and 103.9; $P = 0.93$). These results suggest that the rs174575 moderation of breastfeeding effects on IQ involves genetic differences in children's LC-PUFA metabolism rather than rs174575 differences among lactating women in their milk composition (see SI Table 3).

Discussion

These results suggest that genetic variation in fatty acid metabolism moderates breastfeeding effects on children's cognitive development. We confirmed the gene–environment interaction in two independent, well characterized birth cohorts who were assessed with the best available intelligence tests, and the combined interaction effect across the two studies yielded a P value of 0.005. Among GG homozygote children, the IQ advantage associated with breastfeeding was nil (−0.1 IQ points). In contrast, among children who were C-carriers, the IQ advantage was 6.8 IQ points (or 0.48 standard deviation units in the general population). This advantage corresponds to a moderate effect size (31) that is associated with many important life outcomes (32). Moreover, we were able to rule out potential confounding of the gene–environment interaction due to gene–exposure correlation, intrauterine growth differences, social class differences, and maternal cognitive ability. This study is an initial step in the process of constructive replication (33); further investigation to replicate and explain this specific gene–environment interaction is warranted.

The genetic moderation of breastfeeding effects on IQ is unlikely to be directly caused by the analyzed SNP, and the molecular mechanism by which rs174575 may influence cognitive development is not known. However, rs174575 demonstrates LD with several SNPs in the *FADS1 FADS2* gene cluster that have been shown to track fatty acid composition of serum phospholipids in humans (27) (Fig. 2); the rs174575 C allele is linked with the major alleles of these SNPs, which are associated with more efficient fatty acid processing, possibly due to increased transcriptional activity or to a more active protein. It may be that rs174575 influences the biosynthesis of n-6 and n-3 series LC-PUFAs from their dietary precursors. Because n-6 and n-3 fatty acids compete for the same desaturase metabolic enzymes at multiple steps, this process may operate differently in breastfed versus nonbreastfed children, whose n-6:n-3 precursor ratio is different (34). Genetic variants may also condition the feedback regulation of polyunsaturated fatty acids. The availability of dietary DHA (as in breast milk) alters gene expression in n-6 and n-3 pathways as well as expression of genes involved in synaptic plasticity (35, 36), and the genomic structure surrounding rs174575 (Fig. 2) contains sterol regulatory element transcription factor binding sites through which polyunsaturated fatty acids modify gene expression (24). In addition, through these various mechanisms, several classes of eicosanoids may also be affected and influence brain development (37).

Our finding has implications for neuroscience and early child development. Human milk is widely promoted as good for the brain, and DHA and AA may be needed for optimal intellectual development (23). However, although evidence for this process is mounting, it has not yet been proven (12). The replicated

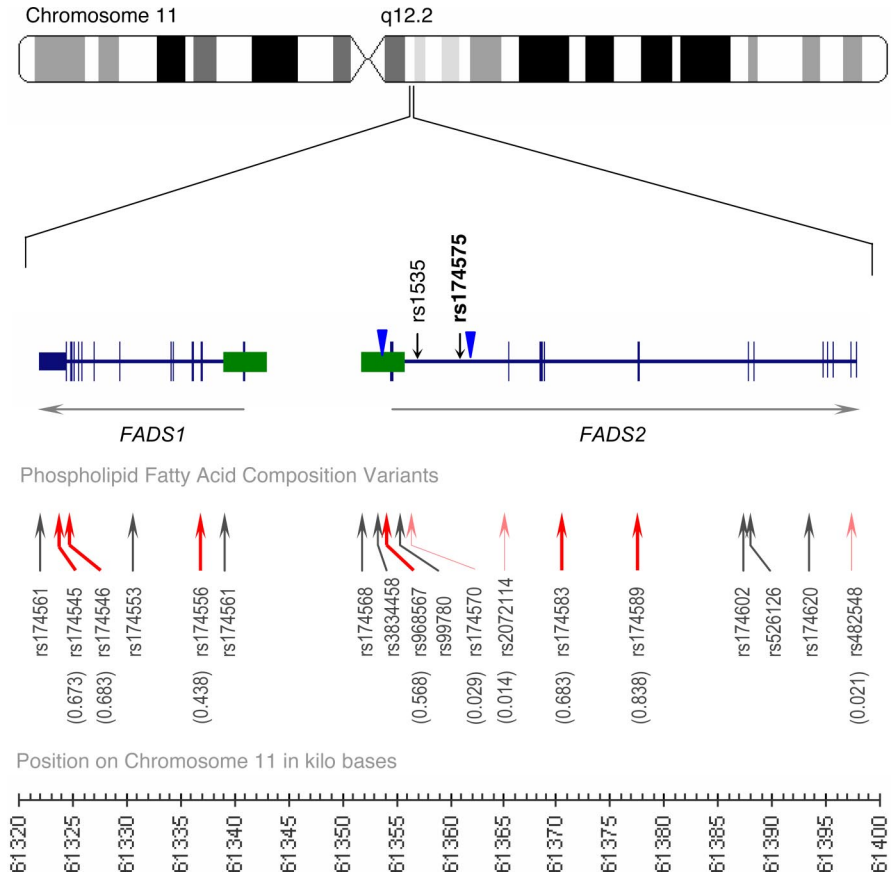


Fig. 2. The *FADS1* and *FADS2* gene structure on chromosome 11q12.2. The *FADS1* and *FAD2* genes are arranged in a head-to-head orientation. The figure shows the position of rs174575 in relation to the hypothetical promoter regions (green rectangles), hypothetical transcription factor binding site SRE (blue arrowheads) (54), and 18 SNPs that have been related to individual differences in fatty acid composition among humans (27). Of these 18 SNPs, 9 (shown in red) are validated in publicly available data via HapMap. r^2 between rs174575 and each of the nine markers found in HapMAP CEU (26) is noted in parentheses. Further suggesting common regulatory mechanisms for these two genes, publicly available gene expression data on lymphoblastoid cell lines from 57 unrelated CEPH individuals (GEO accession no. GDS2106) show that *FADS1* gene expression correlates 0.78 with *FADS2* gene expression. Similarly, in mouse BXD strains, *FADS1* and *FADS2* gene expression correlate 0.61 (55).

finding here that breastfeeding effects on IQ depend on genetic variation in fatty acid metabolism supports a likely neurobiological pathway uniting the gene, environmental exposure, and phenotype and strengthens the need for nutrigenomic studies to trace it in the laboratory (38). Randomized controlled trials comparing the neurodevelopment of infants fed DHA-supplemented versus unsupplemented formula have yielded inconsistent results (15). Our finding suggests an intriguing explanation. Unobserved genetic heterogeneity in fatty acid metabolism—and in response to other supplemented nutrients—may dilute supplementation effects. If genetically responsive subgroups can be identified for analysis, modest benefits may be revealed as stronger than previously thought for some children.

The finding also has implications for behavioral genomics. It is reasonable to ask whether *FADS2* is a "gene for" IQ. There was no overall main effect of genotype on IQ. However, breastfed rs174575 C-carriers scored 4.1 IQ points higher than GG homozygotes (104.0 vs. 99.9, $P = 0.02$). This finding suggests that under human ancestral conditions, when all infants were breastfed, genetic variation in *FADS2* could have influenced individual differences in intelligence. The interaction also offers a clue about why the *FADS2* locus has not appeared in the first genome-wide scans for intelligence (39). Genome-wide scans aim to uncover genes having direct main effects (i.e., genes that show associations with phenotypes regardless of participants' environments). Such scans are inefficient for detecting genes whose effects are conditional on environmental exposures. In contemporary samples of which a nontrivial proportion of participants are not fed breast milk, any link between variation in *FADS2* and IQ would be concealed. The link between genetic variation in fatty acid metabolism and IQ can be revealed only against a specific environmental background (that was universal before formula feeding). This hypothesis suggests that biological interrogation of *FADS2* in relation to intelligence can be pursued further in mammalian species. Comparative genomics shows that rs174575 is in a sequence conserved between human (hg17), chimpanzee (panTro1), dog (canFAm1), rat (rn3), and mouse (mm5) having regulatory potential (40), and there are nonhuman animal models for dietary deprivation/supplementation (41) and for the study of intelligence (42).

Finally, the finding has implications for the public understanding of genetics. The pendulum of opinion surrounding nature versus nurture has swung back and forth, yielding global estimates of heritability versus "environmentality" that have overlooked the contribution of interactions between specific genes and specific experiences (6). To date, research on gene–environment interactions has been dominated by the search for genetic variants that increase disease susceptibility to environmental pathogens (43); for example, carriers of "short" *5-HTT* alleles who encounter stressful life events are at risk of becoming depressed (44); carriers of "rapid" *NAT2* alleles who eat red

meat are at risk of developing colorectal cancer (45). However, genes are not only implicated in disease; here we have shown that a genetic variant (in *FADS2*) may also enhance a favorable response (increased IQ) to a salubrious exposure present throughout human ancestry (breastfeeding).

Materials and Methods

Study 1. Participants in the first cohort were members of the Dunedin Multidisciplinary Health and Development Study, which tracks the health and behavior of a birth cohort of 1,037 children. This sample was constituted at age 3, when the investigators enrolled 91% of consecutive births between April 1972 and March 1973 in Dunedin, New Zealand. Cohort families represent the full range of socioeconomic status in the general population of New Zealand's South Island. Details about the sample are reported elsewhere (46). The sample has been followed to age 32 with 96% retention.

Breastfeeding was assessed in interviews with mothers when the children were 3 years old; 57% of the cohort children were breastfed, consistent with breastfeeding rates in New Zealand in the early 1970s (47). Those infants not breastfed generally received formula feeding prepared from dried cow's milk powder.

IQ was measured at ages 7, 9, 11, and 13 years with the Wechsler Intelligence Scale for Children–Revised (48). IQ scores from the four assessments were combined to form an overall score. The children's IQ scores ranged from 55 to 147 and were normally distributed.

DNA was obtained from 97% of study members as adults. To avoid potential problems of population stratification, DNA from cohort members of Maori origin was excluded from our analyses.

Study 2. Participants in the second cohort were members of the Environmental Risk (E-risk) Longitudinal Twin Study, which tracks the development of a sample of 2,232 children. This E-risk sample was drawn from a 1994–1995 birth register of twins born in England and Wales (49). The E-risk sample was constructed in 1999–2000, when 1,116 families with same-sex 5-year-old twins (93% of those eligible) participated in home-visit assessments, forming the base cohort for the longitudinal E-risk study. Details about the sample are reported elsewhere (50). Because each study family contains two children, all statistical analyses reported here were corrected conservatively for the nonindependence of the twin observations by using tests based on the sandwich or Huber/White variance estimator (51).

Breastfeeding was assessed by a postal questionnaire completed by mothers when the children were 2 years old; 48% of the cohort children were breastfed. Infants not breastfed received formula feeding (before LC-PUFA supplementation became widely available in the U.K.).

IQ was measured at age 5 by using a short form of the Wechsler Preschool and Primary Scale of Intelligence–Revised (52) comprising Vocabulary and Block Design subtests, and by following procedures described by Sattler (ref. 53, table H-6). The children's IQs ranged from 52 to 145 and were normally distributed.

DNA was obtained via buccal swabs from 2,140 children (96% of the sample) and their mothers. Non-Caucasian participants (9%) were excluded from our analyses.

SNP genotyping protocols are summarized in SI Table 4. In both cohorts, genotyping was performed blind to measures of breastfeeding exposure and IQ phenotype.

This work was supported by the National Institute of Mental Health (United States), the Medical Research Council (United Kingdom), and the Health Research Council (New Zealand). A.C. and T.E.M. are Royal Society–Wolfson Research Merit Award holders.

1. Gould SJ (1981) *The Mismeasure of Man* (Norton, New York).
2. Herrnstein RJ, Murray C (1994) *The Bell Curve* (Free Press, New York).
3. Gray JR, Thompson PM (2004) *Nat Rev Neurosci* 5:471–482.
4. Bouchard TJ, McGue M (1981) *Science* 212:1055–1059.
5. Plomin R, DeFries JC, McClearn GE, McGuffin P (2000) *Behavioural Genetics* (Freeman, New York).
6. Ridley M (2003) *Nature via Nurture* (Harper Collins, New York).
7. Caspi A, Moffitt TE (2006) *Nat Rev Neurosci* 7:583–590.
8. American Academy of Pediatrics (2005) *Pediatrics* 115:496–506.
9. Moffitt TE, Caspi A, Rutter M (2005) *Arch Gen Psychiatry* 62:473–481.
10. Institute of Medicine (2004) *Infant Formula*: *Evaluating the Safety of New Ingredients* (National Academies Press, Washington, DC).
11. Sauerwald TU, Demmelmair H, Koletzko B (2001) *Lipids* 36:991–996.
12. Heird WC, Lapillonne A (2005) *Annu Rev Nutr* 25:549–571.
13. Farquharson J, Cockburn F, Patrick WA, Jamieson EC, Logan RW (1992) *Lancet* 340:810–813.
14. Makrides M, Neumann MA, Byard RW, Simmer K, Gibson RA (1994) *Am J Clin Nutr* 60:189–194.
15. McCann JC, Ames BN (2005) *Am J Clin Nutr* 82:281–295.
16. Anderson JW, Johnstone BM, Remley DT (1999) *Am J Clin Nutr* 70:525–535.
17. Mortensen EL, Michaelson KF, Sanders SA, Reinisch JM (2002) *J Am Med Assoc* 287:2365–2371.
18. Catalan J, Moriguchi T, Slotnick B, Murthy M, Greiner RS, Salem N, Jr (2002) *Behav Neurosci* 116:1022–1031.
19. Carrie I, Guesnet P, Bourre JM, Frances H (2000) *Br J Nutr* 83:439–447.
20. Champoux M, Hibbeln JR, Shannon C, Majchrzak S, Soumi SJ, Salem N, Jr, Higley JD (2002) *Pediatr Res* 51:273–281.
21. Takeuchi T, Fukumoto Y, Harada E (2002) *Behav Brain Res* 131:193–203.
22. Lesa GM, Palfreyman M, Hall DH, Clandinin MT, Rudolph C, Jorgensen EM, Schiavo G (2003) *J Cell Sci* 116:4965–4975.
23. Marszalek JR, Lodish HF (2005) *Ann Rev Cell Dev Biol* 21:633–657.
24. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M (2006) *Nucleic Acids Res* 34:D354–D357.
25. Nakamura MT, Nara TY (2004) *Annu Rev Nutr* 24:345–376.
26. The International HapMap Consortium (2003) *Nature* 426:789–796.
27. Schaeffer L, Gohlke H, Muller M, Heid IM, Palmer LJ, Kompauer I, Demmelmair H, Illig T, Koletzko B, Hienrich J (2006) *Hum Mol Genet* 15:1745–1756.
28. Der G, Batty GD, Deary IJ (2006) *Br Med J* 333:945–948.
29. Matte TD, Bresnahan M, Begg MD, Susser E (2001) *Br Med J* 323:310–314.
30. Lawlor DA, Clark H, Davey Smith G, Leon DA (2006) *Pediatrics* 117:e894–e902.
31. Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Earlbaum Associates, Hillsdale, NJ).
32. Gottfredson LS (1997) *Intelligence* 24:79–132.
33. Lykken DT (1968) *Psych Bull* 70:151–159.
34. Yehuda S (2003) *World Rev Nutr Diet* 92:37–56.
35. Rao JS, Ertley RN, DeMar JJC, Rapoport SI, Bazinet RP, Lee H-J (2006) *Mol Psychiatry* 12:151–157.
36. Kitajka K, Puskás LG, Zvara A, Kackler JL, Barceló-Coblijn G, Yeo YK, Tibor F (2002) *Proc Natl Acad Sci USA* 99:2619–2624.
37. Uauy R, Hoffman DR, Peirano R, Birch DG, Birch EE (2001) *Lipids* 36:885–895.
38. Wainwright PE (2002) *Proc Nutr Soc* 61:61–69.
39. Posthuma D, Luciano M, Geus EJ, Wright MJ, Slagboom PE, Montgomery G, Boomsma DI, Martin NG (2005) *Am J Hum Genet* 36:65–76.
40. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, *et al.* (2003) *Nucleic Acids Res* 31:51–54.
41. Fedorova I, Salem N, Jr (2006) *Prostaglandins Leukot Essen Fatty Acids* 75:271–289.
42. Plomin R (2001) *Nat Rev Neurosci* 2:136–141.
43. Hunter DJ (2005) *Nat Rev Genet* 6:287–298.
44. Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington HL, McClay J, Mill J, Martin J, Brathwaite A, *et al.* (2003) *Science* 301:386–389.
45. Chen J, Stampfer MJ, Hough HL, Garcia-Closas M, Willett WC, Hennekents CH, Kelsey KT, Hunter DJ (1998) *Cancer Res* 58:3307–3311.
46. Moffitt TE, Caspi A, Rutter M, Silva P (2001) *Sex Differences in Antisocial Behaviour*: *Conduct Disorder*, *Delinquency*, *and Violence in the Dunedin Longitudinal Study* (Cambridge Univ Press, Cambridge, UK).
47. Sears MR, Greene JM, Willan AR, Taylor DR, Flannery EM, Cowan JA, Herbison GP, Poulton R (2002) *Lancet* 360:901–907.
48. Wechsler D (1974) *Manual for the Wechsler Intelligence Scale for Children–Revised* (Psychological Corporation, New York).

49. Trouton A, Spinath F, Plomin R (2002) *Twin Res* 5:444–448.
50. Moffitt TE, E-Risk Study Team (2002) *J Child Psychol Psychiatry* 43:201–235.
51. Statacorp (2005) Stata Statistical software, Release 8.2 (Stata Corporation, College Station, TX).
52. Wechsler D (1990) *Wechsler Preschool and Primary Scale of Intelligence–Revised* (Psychological Corporation, London).
53. Sattler JM (1992) *Assessment of Children*: *WISC-III and WPPSI-R Supplement* (Jerome M. Sattler, San Diego).
54. Nara TY, He WS, Tang C, Clarke SD, Nakamura MT (2002) *Biochem Biophys Res Commun* 296:111–117.
55. Gatti D, Maki A, Chesler EJ, Kosyk O, Kirova R, Lu L, Manly KF, Qu Y, Williams RW, Perkins A, *et al.* (2007) *Hepatology* 46:548–557.
56. Thurstone TG, Thurstone LL (1973) *The SRA Verbal Form* (Science Research Associates, Chicago).
57. Wilkinson GS (1993) *The Wide Range Achievement Test* (Wide Range, Inc, Wilmington, DE).

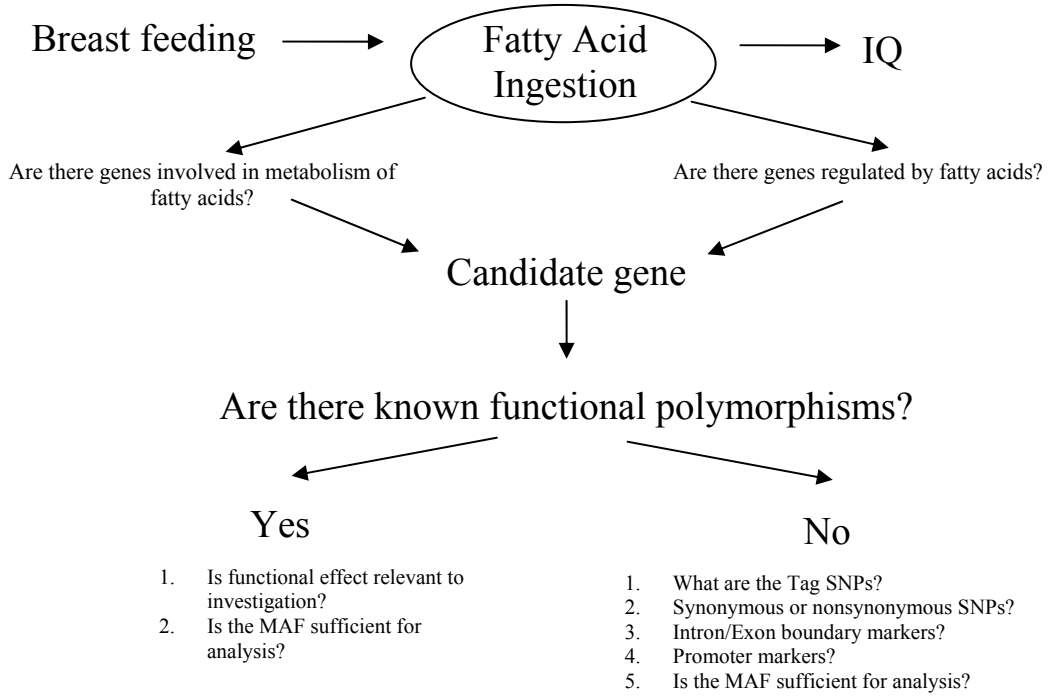
SUPPLEMENTAL INFORMATION TO ACCOMPANY:

Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism

Supplemental Text: Candidate gene and marker selection strategy (to accompany the article: *Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism*).

The flow chart (**Figure 1**) shows the logic of our inquiry in selecting a candidate gene that may moderate the effect of breast feeding on the IQ. It summarizes the process by which we arrived at *FADS2* as our candidate gene and how we selected, a priori, genetic markers for our test of a gene x environment interaction in predicting the IQ.

Figure 1: Flow chart summarizing candidate gene and marker selection strategy

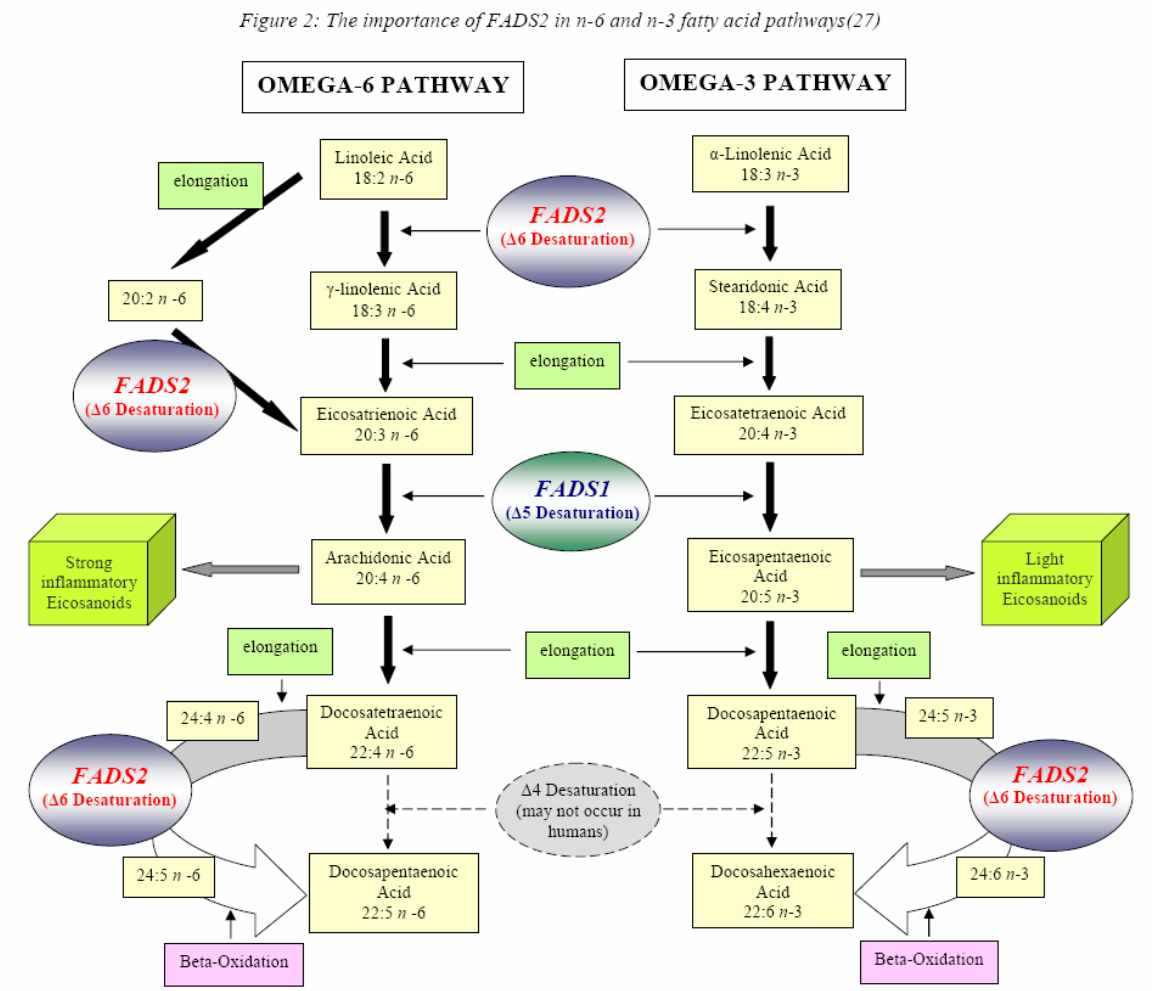


As we have discussed elsewhere (9), the soundest logical basis for selecting a candidate gene for inclusion in the study of a gene-environment interaction is evidence

that the gene is related to the organism's reactivity to the environmental exposure. As such, in the first step of developing our hypothesis, we focused our attention on the putative biological process by which breast feeding is hypothesized to affect IQ. Specifically, breastfeeding is thought to influence brain development via nutritional processes involving fatty acids (10,12,23).

In the second step of developing our hypothesis, we sought to identify candidate genes that are involved in this aforementioned biological pathway. Specifically, we were interested in genes that were involved in the metabolism of fatty acids or in genes whose expression was affected by fatty acids. A review of the literature and the Kegg database (24) led us to *FADS2* as the most obvious candidate gene for two reasons. (a) *FADS2* encodes the delta6 desaturase. As shown in **Figure 2**, delta6 desaturation is a necessary and unavoidable step in the endogenous synthesis of LC-PUFAs from their dietary precursors. Of all the gene products involved in this pathway, delta6 desaturase catalyses the rate limiting step, making it the most crucial element in controlling the availability of endogenously produced AA and DHA. (b) Expression of delta6 desaturase is also regulated via end-product inhibition and dietary LC-PUFAs such as are available in breast milk. *FADS2* was the only gene we tested.

Figure 2: The importance of FADS2 in n-6 and n-3 fatty acid pathways(27)

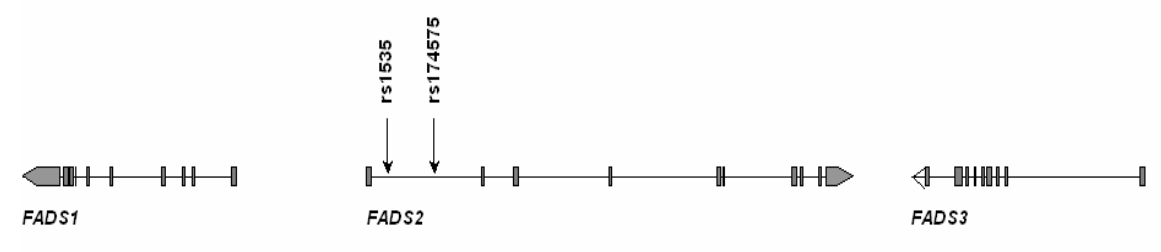


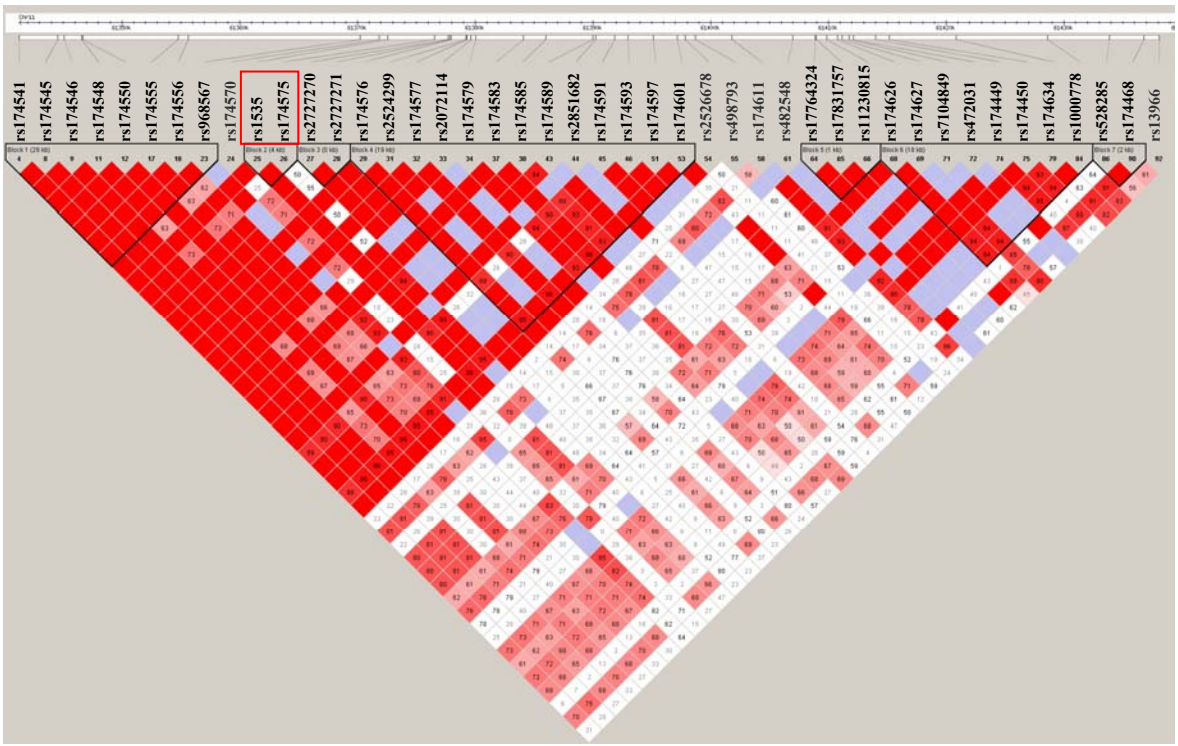
In the third step of developing our hypothesis, we needed to select markers in *FADS2*. To the best of our knowledge, *FADS2* polymorphisms have not been studied in relation to breast milk or to IQ. In this event, a standard approach to marker selection would be to cover the entire gene with Tag SNPs. The difficulty with this approach is that it yields many SNPs with small minor allele frequencies (MAF) that are not useable

in regression analyses of gene-environment interaction as they would lead to statistical tests that would be woefully underpowered (i.e., the statistical power to detect interactions is determined, in tandem, by allele frequency and environmental exposure frequency). For this reason, we used HapMap data (version 20) to strategically select two Tag SNPs for two reasons (as shown in the LD map in **Figure 3** these SNPs, in the red box, are rs1535 and rs174575). (a) Compared to all the Tag SNPs, these two SNPs had minor allele frequencies (.37 and .28, respectively) which ensured that we would have sufficient statistical power to test a hypothesis of a gene x environment interaction and (b) these two SNPs are located to cover variation in substantial parts (including most significantly in the promoter region) of *FADS2* and also of *FADS1*, a gene which lies in a head-to-head orientation with *FADS2* and which is also involved in fatty acid metabolism, encoding the delta5 desaturase (see Figure 2). rs1535 and rs174575 were the only markers we tested.

To summarize, our hypothesis-driven research about gene x environment interaction proceeded in the following fashion: (1) Start with the environmental exposure; (2) specify the putative biological pathway by which it affects the phenotype; (3) identify a candidate gene involved in this same pathway; (3) select the most promising markers; (4) test and replicate.

Figure 3: FADS2 Linkage Disequilibrium (LD) map. Physical location and linkage disequilibrium pattern of SNPs covering the FADS2 gene from the promoter region to the 3' end of the gene (and also including FADS1). The top part of the panel shows the head-to-head orientation of FADS2 and FADS1, and the position of these SNPs. The bottom part of the panel shows an LD plot generated with Haploview, using r^2 and D' as the measure of LD.





Supplemental Table 1. Interpreting the rs174575-breastfeeding interaction: Tests to rule out confounding by social class and maternal cognitive ability.

Testing the possibility of a social-class confound: One way to examine if the breastfeeding measured in this research is a mere proxy for social class is to test whether an interaction between rs174575 and social class also predicts IQ; if the rs174575-breastfeeding interaction was entirely due to the fact that breastfeeding is correlated with social class, we would expect to find a significant interaction between rs174575 and social class. In contrast to the significant rs174575-breastfeeding interaction (shown in Table 1 and Figure 1), this Supplemental Table 2 shows that the rs174575-social class interaction was not significant in either Dunedin (P=.32) or E-risk (P=.64).

Samples	rs174575 CC homozygotes			rs174575 CG heterozygotes			rs174575 GG homozygotes		
	Low class	Medium class	High class	Low class	Medium class	High class	Low class	Middle class	High class
New Zealand (Dunedin) birth cohort	**n=87**	**n=255**	**n=79**	**n=68**	**n=229**	**n=57**	**n=17**	**n=56**	**n=9**
Intelligence quotient (IQ)	93.5 (13.6)	100.5 (13.7)	111.4 (12.8)	92.3 (12.9)	100.0 (11.8)	111.4 (13.6)	88.7 (10.2)	100.9 (10.7)	112.7 (13.5)
British (E-risk study) birth cohort	**n=370**	**n=346**	**n=334**	**n=244**	**n=254**	**n=237**	**n=34**	**n=49**	**n=45**
Intelligence quotient (IQ)	94.3 (13.5)	99 .0 (13.6)	108.6 (13.7)	93.7 (13.7)	99.8 (13.2)	108.6 (14.2)	92.6 (12.9)	98 (16.1)	108.3 (14.8)

Entries in the table are means and standard deviations.

Yet another way to evaluate the results is to compare the correlations between social class and IQ to the correlations between breastfeeding and IQ, in each of the genotype groups. In Dunedin, the correlation between social class and IQ is similar in each of the genotype groups (.39 among C-carriers and .53 among GG homozygotes). In contrast, the correlation between breastfeeding and IQ is .23 versus -.05 in C-carriers and GG homozygotes, respectively. Likewise, in E-risk, the correlation between social class and IQ is similar in each of the genotype groups (.40 among C-carriers and .40 among GG homozygotes). In contrast, the correlation between breastfeeding and IQ is .24 versus .02 in C-carriers and GG homozygotes, respectively.

Supplemental Table 1 (cont.)

Testing the possibility of a maternal cognitive ability confound: The same analysis can be extended to examine if the breastfeeding measured in this research is a mere proxy for maternal cognitive ability. If the rs174575-breastfeeding interaction was entirely due to the fact that breastfeeding is correlated with maternal cognitive ability, we would expect to find a significant interaction between rs174575 and maternal cognitive ability. However, this interaction was not significant in either cohort (P=.26 and P=.98). In Dunedin, the correlation between maternal cognitive ability and child IQ was significant and of similar magnitude in both genotype groups (.41 among C-carriers and .28 among GG homozygotes). In E-risk the correlation between maternal cognitive ability and child IQ was also significant and of similar magnitude in both genotype groups (.31 among C-carriers and .28 among GG homozygotes).

Supplemental Table 2. Interpreting the rs174575-breastfeeding interaction: Tests to rule out maternal genotype effect.

Data in this table can be used to test whether the rs174575 moderation of breastfeeding effects on IQ represents genetic differences in children's fatty acid metabolism (a child-genotype effect) or genetic differences in breastmilk quality (a maternal- genotype effect). To test the maternal genotype effect, we genotyped rs174575 among mothers of children enrolled in the E-risk study. We then derived three testable predictions. First, if what we are observing is a maternal genotype effect, we should find that among breastfeeding mothers, C-carrying mothers have children with higher IQs than GG mothers. They did not; IQ scores of breastfed children of CC, CG, and GG mothers were 104.6, 103.4, and 103.9 (P=.93). Second, we should find that CG children with C-carrying mothers benefit more from breastmilk than CG children with GG mothers. They did not; among CG children, the significant association between breastfeeding and IQ did not vary by maternal genotype (P=.86): Specifically, breastfed-CG children of C-carrying mothers had higher IQs than their non-breastfed counterparts (104.7 vs. 97.0) and breastfed-CG children of GG mothers also had higher IQs than their non-breastfed counterparts (106.8 vs. 99.9). Third, we should find that GG children with CG mothers benefit more from breastmilk than GG children with GG mothers. They did not; among GG children, the association between breastfeeding and IQ did not vary by maternal genotype (P=.81). Specifically, there was no breastfeeding effect among GG children of CG mothers (100.5 vs. 99.6) and no breastfeeding effect among GG children of GG mothers (98.1 vs. 99.0).

	Children's rs174575 genotype													
	CC				CG						GG			
Mothers' rs174575 genotype	CC		CG		CC		CG		GG		CG		GG	
	Not breastfed	Breast-fed	Not breastfed	Breast-fed	Not breastfed	Breast-fed	Not breastfed	Breast-fed	Not breastfed	Breast-fed	Not breastfed	Breast-fed	Not breastfed	Breast-fed
N	361	353	116	109	137	141	166	142	49	38	37	43	21	18
IQ (SD)	97.7 (13.9)	104.3 (15.3)	96.9 (15.4)	104.1 (13.9)	97.1 (13.6)	105.6 (14.8)	97.0 (13.4)	103.9 (15.4)	99.9 (17.1)	106.8 (13.8)	99.6 (16.1)	100.5 (16.8)	99.0 (14.7)	98.1 (17.5)

Supplemental Table 3. Genotyping details for two common variants in the *FADS2* gene.

In both cohorts, the two *FADS2* polymorphisms were genotyped using manufacturer recommended protocols on the AB7900 TaqMan platform. The following functionally tested, made-to-order SNP genotyping assays from Appliedbiosystems (56) were used (the assay IDs in the table link directly to the Appliedbiosystems web page for ordering):

rs174575	C___2575522_20
rs1535	C___2575527_10

The distributions of genotypes in both cohorts were consistent with Hardy-Weinberg equilibrium (HWE). D' = . 94 and .95 between the two SNPs in each of the two cohorts (r^2 = .77 and .79, respectively).

dbSNP	Alleles (major/minor)	Allele frequency	HWE (P value)
rs174575	C/G	.70/.30 (Dunedin) .74/.26 (E-risk)	.47 (Dunedin) .84 (E-risk)
rs1535	A/G	.61/.39 (Dunedin) .66/.34 (E-risk)	.94 (Dunedin) .76 (E-risk)

To test the maternal genotype effect, we used the same procedure to genotype rs174575 among E-risk mothers.

dbSNP	Alleles (major/minor)	Allele frequency	HWE (P value)
rs174575	C/G	.75/.25 (E-risk)	.22 (E-risk)