Genetic Approaches to Studying Common Diseases and Complex Traits

JOEL N. HIRSCHHORN

*Program in Genomics and Division of Endocrinology, Children's Hospital, Boston, Massachusetts, 02115,
Departments of Genetics and Pediatrics, Harvard Medical School, Boston, Massachusetts, 02115,
Metabolism Initiative and Program in Medical and Population Genetics, Broad Institute of MIT and
Harvard, Cambridge, Massachusetts, 02139*

ABSTRACT

Most common diseases and most quantitative traits that can be measured in human populations are complex genetic traits. That is, many genetic and nongenetic factors interact to determine the final phenotype, whether that phenotype is susceptibility to disease, or a quantifiable trait such as height, weight, serum cholesterol, or blood pressure. Identifying the genes that underlie the population variation in these phenotypes has been challenging. Recently, databases of common genetic variants, recognition of the patterns of genetic variation, and rapid genotyping methodologies have emerged, and the combination of these tools and

resources will greatly facilitate genetic association studies, a potentially powerful method to map the genes for complex traits. However, care will be required in performing and interpreting these association studies. Until genome-wide studies are feasible, choosing candidate genes will be necessary. In addition, the choice of phenotype will likely influence the success of these gene mapping efforts. Finally, population genetic methods, including searching for genes under selection, may provide clues to the location of the genes for common disease and complex traits. (*Pediatr Res* 57: 74R–77R, 2005)

GENETIC BASIS OF COMPLEX TRAITS AND COMMON DISEASE

Many common diseases cluster in families in patterns that demonstrate that genetics plays a role in determining susceptibility. For example, the identical twin of a patient with type 1 diabetes will also get type 1 diabetes 30–50% of the time; dizygotic twins (who share a common environment but only 50% of their genes) are much less concordant (1,2). A sibling of a patient with type 1 diabetes is 15 times more likely to get diabetes than an unrelated individual (3), also suggesting a strong genetic component to disease susceptibility. In the case of schizophrenia, Risch (4) further demonstrated that the risk of disease falls off rapidly for relatives of schizophrenic individuals with decreasing genetic relatedness, consistent with a model where variation in multiple genes combines to influence disease risk. The increased risk to relatives (λ) is one measure of the influence of genetics; another measure of the contribution of inherited factors is termed heritability (h^2), which

signifies the fraction of the population variation that can be explained by genetic factors working together in an additive fashion. Heritability can be estimated either from family studies or twin studies, and is often between 30 and 50% for common diseases such as diabetes (1,2,5), or quantitative traits such as body mass index or blood pressure (6,7). Thus, multiple genetic factors and nongenetic factors combine to influence the risk of common diseases and quantitative traits. Because multiple genetic and nongenetic factors interact to affect phenotype, these diseases and traits are termed complex genetic traits.

For most common diseases and complex traits, the underlying genetic variation remains unknown. In principle, the relevant genetic variation could be rare (alleles with frequencies well under 1% in the population), as is true for most single-gene disorders, or more common (frequencies above 1% in the population). The frequency spectrum of the alleles for complex traits is important to consider, because it will guide approaches to finding the causal genetic variants (see below). Although the allele frequencies are largely unknown (because the causal variants remain largely unknown), theoretical and empirical considerations suggest that for common diseases and complex traits, some of the causal genetic variants may be common (8–10). In particular, most of the genetic variation that one encounters in the human population is explained by common variants with allele frequencies of 5% or greater (11,12). Because the bulk of random genetic variation is presumably

Received January 5, 2005; accepted January 5, 2005.

Correspondence: Joel N. Hirschhorn, M.D., Ph.D., Children's Hospital, Boston, Department of Pediatrics, 300 Longwood Ave., Enders 651, Boston, MA 02115; e-mail: joel.hirschhorn@childrens.harvard.edu

J.N.H. is a recipient of a Smith Family Foundation New Investigator Award, and a Burroughs Wellcome Career Award in Biomedical Sciences.

J.N.H. was the recipient of the Society for Pediatric Research 2004 Young Investigator Award presented at the 2004 Annual Meeting of the Pediatric Academic Societies, San Francisco, CA.

evolutionarily neutral (neither under strong positive or negative selection), this suggests suggesting that disease variants that are not strongly evolutionarily deleterious will likewise be common in the population (10,13). But, we already know for most single-gene disorders that the responsible alleles are generally quite rare [unless there is a balancing selective pressure, such as malaria resistance for sickle cell disease (14)]. Why should the variants that cause common disease or affect quantitative traits not also be predominantly rare?

There are at least four arguments in favor of a role for common variation in complex traits. First, common diseases (or high or low quantitative trait values) are generally not as evolutionarily disadvantageous as single-gene disorders, which often cause early death or at least markedly decreased reproductive capability. Second, the variants that cause single-gene disorders are highly penetrant (meaning each variant is sufficient, or nearly so, to cause disease), whereas multiple variants are required to cause common diseases or strongly influence quantitative traits. Thus, the impact of selective pressure is diluted for the variants for complex traits. Third, most single-gene diseases are rare, whereas most polygenic diseases are common; population genetic arguments predict that for common diseases, some of the causal genetic variation should have a high frequency in the population, due to the demographic history of the human population (10). Similar arguments can be made for quantitative traits that vary across the entire population. Finally, empirical evidence suggests that common variants do contribute to the risk of common diseases (15).

APPROACHES TO FINDING THE GENES FOR COMPLEX TRAITS

Two major approaches have been used to map genetic variants that influence disease risk: linkage analysis and association studies. In linkage analysis, a genome-wide set of a few hundred or a few thousand markers spaced millions of bases apart is typed in families with multiple affected relatives (or multiple relatives in whom a trait has been measured). Markers that segregate with disease (or the trait) in relatives more often than expected are used to localize the disease genes. This approach has the advantage of being an unbiased, comprehensive search across the genome for susceptibility alleles, and has been successfully applied to find the genes for many single-gene disorders. However, linkage analysis has been less successful for polygenic diseases and quantitative traits [(16); see (17) for discussion], perhaps in part because of a limited power to detect the effect of common alleles with modest effects on disease (18,19).

Association studies look for a particular marker to be correlated with disease (or trait values) across a population rather than within families. These studies have much greater power to detect the effects of common variants (4). For example, the insulin VNTR class III allele, which has a frequency of approximately 70%, has definitively shown to modestly affect the risk of type 1 diabetes, with a p value of 10^{-22} , using association studies (20). By contrast, the region containing the insulin gene is just barely above the threshold for statistical significance even when all of the world's linkage data for type 1 diabetes is combined (21). Similarly, the common Pro12Ala

polymorphism reproducibly affects the risk of type 2 diabetes, but studies of millions of sib pairs would be required to get a significant signal using linkage (5,19).

However, association studies require many more markers than linkage analysis. In linkage analysis, the markers must merely be in linkage with the disease allele (that is, the marker and the disease allele must generally be inherited together within the one or two generations spanned by a family). Thus, the markers can be several million bases away from the relevant gene. By contrast, association studies require that the markers be in linkage disequilibrium with the disease allele (*i.e.* the marker and the disease allele must generally be inherited together throughout the population, despite the many thousands of generations that have usually elapsed between a population's common ancestor and the present day). Because segments of linkage disequilibrium are measured in tens of thousands of bases (rather than the tens of millions of bases for linkage), hundreds of thousands of markers will be required to scan the genome for association (11,22). In addition, association studies, although potentially more powerful than linkage, still require samples sizes of thousands of individuals. Thus, association studies are currently limited to candidate genes or regions, because of the expense of a genome-wide approach.

INTERPRETING ASSOCIATION STUDIES

Association studies of complex traits also present an additional challenge, in that they are often misinterpreted and therefore appear to be poorly reproducible. If there is a "significant" correlation between genotypes at a variant and a disease or quantitative trait, that is, the p value is 0.05 or lower, this degree of correlation is usually interpreted as evidence of association between the variant and the phenotype. However, most such associations are not consistently reproduced—in a review of the literature, only 6 of 166 associations were replicated by at least 75% of subsequent studies (23). The possible causes of this inconsistency are false-positive reports, false-negative studies that incorrectly fail to replicate a valid association, or true heterogeneity between studies. We performed a meta-analysis of association studies (15), and found that of 25 reported associations, over half had no evidence of replication, indicating that the false-positive rate in the literature is high (using the $p < 0.05$ standard). However, a sizable fraction [8/25 associations in our study, and a comparable fraction in a similar study (24)] showed evidence of replication. In these cases, the failure to consistently replicate the association was likely explained by the modest effects of the causal variants on disease risk and the consequent false-negative studies. In most cases, the causal variant was associated with a 10–50% increased risk of disease, meaning that sample sizes in the thousands are required to achieve even a nominally significant p value < 0.05 . Because most association studies had used samples of hundreds of individuals, the lack of consistency even for *bona fide* causal variants is not surprising.

An illustration of the importance of large sample sizes can be seen in the association between the *PPARG* Pro12Ala polymorphism and type 2 diabetes. This missense variant (25) was first reported to have a 3-fold effect on diabetes risk, with

the more common proline-encoding allele conferring higher risk (26). Four of five subsequent studies reported that there was no association (27–31), because the evidence for association did not reach the nominal significance level of $p < 0.05$. However, our larger study both confirmed the association and demonstrated that the effect on diabetes risk was more modest than originally described [about a 25% increased risk associated with the proline variant (19)]. The association has since been confirmed in several other large studies (32–34), reaching an overall p value of less than 10^{-9} after more than 20,000 patients were examined (Fig. 1). The modest effect of this variant on diabetes risk likely explains most of the negative studies, which predominantly trended in the same direction and were mostly consistent with the overall estimate of the effect of this variant. Also illustrated by these data are a common phenomenon in association studies, the “winner’s curse,” in which the first report overestimates the genetic effect size (15,24)—the first reported association between the Pro12Ala variant and type 2 diabetes (26) provided the largest estimate of the effect of this variant on diabetes risk (Fig. 1).

CURRENT AND FUTURE DIRECTIONS

In the future, it will become possible to screen the vast majority of common genetic variation for a role in complex traits. These studies will be greatly facilitated by the recognition of that many common variants are strongly correlated (in

linkage disequilibrium), and hence redundant (11), meaning that a few hundred thousand single nucleotide polymorphisms (SNPs) will suffice to survey the approximately ten million variants with frequency 5% or greater (17,22). However, such studies are still too laborious and expensive for routine application to complex traits. Until they are more practical, we must still select genes and variants to study, and the choice of trait will also be important to maximize the chance of success.

We have chosen to study both common diseases (because of their direct medical relevance) and quantitative traits (because they are easier to measure accurately in large populations, often have a stronger genetic component, and are often themselves risk factors for disease). For example, we study stature (adult height) as a model quantitative trait because of its ease of measurement and high heritability. We have successfully performed linkage analysis to identify several regions that likely harbor genes that affect adult height (35). Several of these have been confirmed (36–38), and we will follow up these findings with association studies of genes in these regions of linkage. We are also focusing on studies of body mass index, a measure of obesity that predicts future diabetes, heart disease, and mortality.

To increase the likelihood that a candidate gene will be harbor variation relevant to traits we are studying, we integrate a wide variety of information, including previous genetic studies, expression analysis, linkage data, animal models, and knowledge of biologic pathways. We also aim within each gene to both broadly survey the majority of common variation but also capture any putative functional variants such as missense polymorphisms, and potentially also variation in evolutionarily conserved noncoding regions (39). Finally, genetic variants that have been under evolutionary selection are by definition functional and thus may be more likely to contribute to disease susceptibility (*e.g.* the sickle cell variant has been selected for its protective effect against malaria). To find such variants, we need to be able to recognize the genetic signatures of selection.

To help characterize the genetic signatures of selection, we studied variation around the gene encoding lactase, *LCT*. Most people and all other mammals lose the ability to digest lactose, which comes exclusively from dairy (40,41). However, certain human populations retain this ability into adulthood, and this trait is inherited in association with a genetic variant upstream of the lactase gene (42). By studying the length of haplotypes at this locus and the frequencies of genetic variants near *LCT* in different populations, we provided the first population genetics-based evidence that there has been strong selective pressure in favor of lactase persistence in Northern European populations. Indeed, the signature of selection we described is one of the strongest in the entire human genome. Genome-wide distributions for haplotype length and allele frequencies in different populations can now be calculated from genotype data in public databases such as the International HapMap Project (www.hapmap.org) or dbSNP (www.ncbi.nlm.nih.gov/SNP). Using these distributions, we plan to search the remainder of the genome for additional signatures of selection.

Figure 1. The association between *PPARG* Pro12Ala and type 2 diabetes becomes more apparent as the sample size increases. Each line represents a different study; the point represents the point estimate of the odds ratio (OR) for the alanine allele (OR <1 suggest a protective effect of the alanine allele on diabetes risk), and the line represents the 95% confidence interval (CI) around the point estimate. The studies are arranged by increasing number of alanine alleles in the study, which is indicative of power. The combined data, representing more than 5,000 alanine alleles in over 20,000 individuals, is shown on the last line, and corresponds to an OR of 0.82 (95% CI, 0.77–0.87), with a p value of 1.8×10^{-10} . That is, carriage of an alanine allele is associated with an approximately 20% decreased risk of type 2 diabetes. The initial study that reported an association between Pro12Ala and type 2 diabetes (26) is indicated with the arrow, and reports the largest observed effect of this variant on diabetes risk, consistent with the winner’s curse (15).

SUMMARY

Understanding the genetic basis of complex traits and common disease is an important goal, since the pathways that affect the risk of disease in patients are also potentially good drug targets. For example, two of the alleles known to affect the risk of type 2 diabetes are in genes (*PPARG* and *KCNJ11*) that also encode drug targets (5). Knowledge of the relevant variants may also aid in prediction and prognosis and in designing optimal therapies and intervention. Identifying the causal genetic variants will likely require a number of different approaches, but association studies are likely to be a significant tool in the near future. However, care must be taken in performing and interpreting these studies, so as to avoid generating false leads and also to correctly identify the *bona fide* genetic risk factors for disease.

Acknowledgments. The author thanks the members of his laboratory, numerous collaborators, and study participants, without whom the research described here would not be possible.

REFERENCES

- Hirschhorn JN 2003 Genetic epidemiology of type 1 diabetes. *Pediatr Diabetes* 4:87–100
- Redondo MJ, Fain PR, Eisenbarth GS 2001 Genetics of type 1A diabetes. *Recent Prog Horm Res* 56:69–89
- Spielman RS, Baker L, Zmijewski CM 1980 Gene dosage and susceptibility to insulin-dependent diabetes. *Ann Hum Genet* 44:135–150
- Risch N 1990 Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 46:222–228
- Florez JC, Hirschhorn J, Altshuler D 2003 The inherited basis of diabetes mellitus: implications for the genetic analysis of complex traits. *Annu Rev Genomics Hum Genet* 4:257–291
- Atwood LD, Heard-Costa NL, Cupples LA, Jaquish CE, Wilson PW, D'Agostino RB 2002 Genomewide linkage analysis of body mass index across 28 years of the Framingham Heart Study. *Am J Hum Genet* 71:1044–1050
- Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavvas H, Cupples LA, Myers RH 2000 Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham Heart Study. *Hypertension* 36:477–483
- Lander ES, Schork NJ 1994 Genetic dissection of complex traits. *Science* 265:2037–2048
- Chakravarti A 1999 Population genetics—making sense out of sequence. *Nat Genet* 21:56–60
- Reich DE, Lander ES 2001 On the allelic spectrum of human disease. *Trends Genet* 17:502–510
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D 2002 The structure of haplotype blocks in the human genome. *Science* 296:2225–2229
- Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D; International SNP Map Working Group 2001 A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409:928–933
- Pritchard JK 2001 Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet* 69:124–137
- Li WH 1975 The first arrival time and mean age of a deleterious mutant gene in a finite population. *Am J Hum Genet* 27:274–286
- Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN 2003 Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 33:177–182
- Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M 2001 Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 69:936–950
- Hirschhorn JN, Daly MJ 2005 Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6:95–108
- Risch N, Merikangas K 1996 The future of genetic studies of complex human diseases. *Science* 273:1516–1517
- Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, Schaffner SF, Bolck S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES 2000 The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26:76–80
- Barratt BJ, Payne F, Lowe CE, Hermann R, Healy BC, Harold D, Concannon P, Gharani N, McCarthy MI, Olavesen MG, McCormack R, Guja C, Ionescu-Tirgoviste C, Undlien DE, Ronningen KS, Gillespie KM, Tuomilehto-Wolf E, Tuomilehto J, Bennett ST, Clayton DG, Cordell HJ, Todd JA 2004 Remapping the insulin gene/IDDM2 locus in type 1 diabetes. *Diabetes* 53:1884–1889
- Cox NJ, Wapelhorst B, Morrison VA, Johnson L, Pinchuk L, Spielman RS, Todd JA, Concannon P 2001 Seven regions of the genome show evidence of linkage to type 1 diabetes in a consensus analysis of 767 multiplex families. *Am J Hum Genet* 69:820–830
- Carlson CS, Eberle MA, Kruglyak L, Nickerson DA 2004 Mapping complex disease loci in whole-genome association studies. *Nature* 429:446–452
- Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K 2002 A comprehensive review of genetic association studies. *Genet Med* 4:45–61
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG 2001 Replication validity of genetic association studies. *Nat Genet* 29:306–309
- Yen CJ, Beamer BA, Negri C, Silver K, Brown KA, Yarnall DP, Burns DK, Roth J, Shuldiner AR 1997 Molecular scanning of the human peroxisome proliferator-activated receptor gamma (hPPAR gamma) gene in diabetic Caucasians: identification of a Pro12Ala PPAR gamma 2 missense mutation. *Biochem Biophys Res Commun* 241:270–274
- Deeb SS, Fajas L, Nemoto M, Pihlajamaki J, Mykkanen L, Kuusisto J, Laakso M, Fujimoto W, Auwerx J 1998 A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nat Genet* 20:284–287
- Hara K, Okada T, Tobe K, Yasuda K, Mori Y, Kadowaki H, Hagura R, Akanuma Y, Kimura S, Ito C, Kadowaki T 2000 The Pro12Ala polymorphism in PPAR gamma2 may confer resistance to type 2 diabetes. *Biochem Biophys Res Commun* 271:212–216
- Mancini FP, Vaccaro O, Sabatino L, Tufano A, Rivellese AA, Riccardi G, Colantuoni V 1999 Pro12Ala substitution in the peroxisome proliferator-activated receptor-gamma2 is not associated with type 2 diabetes. *Diabetes* 48:1466–1468
- Ringel J, Engeli S, Distler A, Sharma AM 1999 Pro12Ala missense mutation of the peroxisome proliferator-activated receptor gamma and diabetes mellitus. *Biochem Biophys Res Commun* 254:450–453
- Clement K, Hercberg S, Passinge B, Galan P, Varroux-Vial M, Shuldiner AR, Beamer BA, Charpentier G, Guy-Grand B, Froguel P, Vaisse C 2000 The Pro115Gln and Pro12Ala PPAR gamma gene mutations in obesity and type 2 diabetes. *Int J Obes Relat Metab Disord* 24:391–393
- Meirhaeghe A, Fajas L, Helbecque N, Cotel D, Auwerx J, Deeb SS, Amouyel P 2000 Impact of the peroxisome proliferator-activated receptor gamma2 Pro12Ala polymorphism on adiposity, lipids and non-insulin-dependent diabetes mellitus. *Int J Obes Relat Metab Disord* 24:195–199
- Douglas JA, Erdos MR, Watanabe RM, Braun A, Johnston CL, Oeth P, Mohlke KL, Valle TT, Ehnholm C, Buchanan TA, Bergman RN, Collins FS, Boehnke M, Tuomilehto J 2001 The peroxisome proliferator-activated receptor-gamma2 Pro12Ala variant: association with type 2 diabetes and trait differences. *Diabetes* 50:886–890
- Mori H, Ikegami H, Kawaguchi Y, Seino S, Yokoi N, Takeda J, Inoue I, Seino Y, Yasuda K, Hanafusa T, Yamagata K, Awata T, Kadowaki T, Hara K, Yamada N, Gotoda T, Iwasaki N, Iwamoto Y, Sanke T, Nanjo K, Oka Y, Matsutani A, Maeda E, Kasuga M 2001 The Pro12 → Ala substitution in PPAR-gamma is associated with resistance to development of diabetes in the general population: possible involvement in impairment of insulin secretion in individuals with type 2 diabetes. *Diabetes* 50:891–894
- Ardlie KG, Lunetta KL, Seielstad M 2002 Testing for population subdivision and association in four case-control studies. *Am J Hum Genet* 71:304–311
- Hirschhorn JN, Lindgren CM, Daly MJ, Kirby A, Schaffner SF, Burt RP, Altshuler D, Parker A, Rioux JD, Platko J, Gaudet D, Hudson TJ, Groop LC, Lander ES 2001 Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *Am J Hum Genet* 69:106–116
- Perola M, Ohman M, Hiekkalinna T, Leppavuori J, Pajukanta P, Wessman M, Koskenvuo M, Palotie A, Lange K, Kaprio J, Peltonen L 2001 Quantitative-trait-locus analysis of body-mass index and of stature, by combined analysis of genome scans of five Finnish study groups. *Am J Hum Genet* 69:117–123
- Xu J, Bleecker ER, Jongepier H, Howard TD, Koppelman GH, Postma DS, Meyers DA 2002 Major recessive gene(s) with considerable residual polygenic effect regulating adult height: confirmation of genomewide scan results for chromosomes 6, 9, and 12. *Am J Hum Genet* 71:646–650
- Wu X, Cooper RS, Boerwinkle E, Turner ST, Hunt S, Myers R, Olshen RA, Curb D, Zhu X, Kan D, Luke A 2003 Combined analysis of genomewide scans for adult height: results from the NHLBI Family Blood Pressure Program. *Eur J Hum Genet* 11:271–274
- Pennacchio LA, Rubin EM 2001 Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet* 2:100–109
- Simoons FJ 1969 Primary adult lactose intolerance and the milking habit: a problem in biologic and cultural interrelations. I. Review of the medical research. *Am J Dig Dis* 14:819–836
- Simoons FJ 1970 Primary adult lactose intolerance and the milking habit: a problem in biologic and cultural interrelations. II. A culture historical hypothesis. *Am J Dig Dis* 15:695–710
- Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Jarvela I 2002 Identification of a variant associated with adult-type hypolactasia. *Nat Genet* 30:233–237