

COMMENT

# Hypothesis-driven genomics pays off

Gregory A Petsko\*

In human genetics, the equivalent of finding gold mine is finding a genetically isolated population with excellent medical records. Inbreeding often magnifies the effects of single gene traits because homozygosity is more common, and a society in which family medical histories can be traced back for many generations allows patterns of inheritance to be ascertained with confidence. Given the mobility of the human population today, such societies have become rare, but they do exist. There was a lot of excitement in the early 1990s about genetic studies of Mormon communities in Utah in the USA, and they have provided useful information. So have French Canadian communities in eastern Canada and some groups in rural India. But the society that attracted the most attention - and the program that has created the most controversy - is that of Iceland.

Geographically isolated for millenia, and possessing a manageably small (around 300,000), highly educated, homogeneous population with a strong tradition of maintaining family histories - genealogical records go back to the first census, in 1703, for most families, and some stretch for over 11 centuries - Iceland would seem a perfect place for finding disease-associated genes. And that's exactly what Dr Kári Stefánsson, an Icelandic-born neurologist who trained in the USA (he was professor of neuropathology at Beth Israel Hospital in Boston in the mid 1990s) thought when, in 1996, he founded deCode genetics. Stefánsson was convinced that he could use the genetic information from his fellow countrymen to identify genes for complex disorders such as schizophrenia, multiple sclerosis and cardiovascular disease. The original plan was to create a genealogical database, collect biobank specimens for genotyping, and establish a national electronic health record system to connect the genetic data to each individual's phenotype. Apart from privacy issues, the plan attracted a lot of negative publicity because deCode was founded as a for-profit company, and the information it collected was intended to generate revenue through, for example, development of diagnostic tests.

In 1998 deCode successfully lobbied the Parliament of Iceland to pass a law establishing a bidding process for the right to create and use an Icelandic Health Sector Database containing the medical records, genealogical and genetic data of all Icelanders. The bidding was won by - surprise - deCode. A decision by the Supreme Court of Iceland in November 2003 effectively killed off the Health Sector Database project (in large part due to concerns about privacy issues), but the company has continued to try to identify disease-related genes without it.

Like most biotechnology companies, deCode's road to profitability has been rocky. It declared bankruptcy in 2009, but has remained in business under the protection of so-called Chapter 11 laws. To create a revenue stream, it has offered "personal genotyping". For around US\$1000, you can send a cheek swab to deCode, and the company will scan your DNA for more than 1,000,000 single nucleotide polymorphisms, some of which are associated with disease risk and/or ancestry.

Despite its financial woes, deCode has begun to pay off, not in money, but in science. About 10 years ago its scientists co-discovered the association of the neuroregulin 1 (*NRG1*) gene with schizophrenia, and they have had some more modest successes with other diseases. But last week, deCode produced not only its most important genetic finding to date, but also quite possibly the most important genetic finding in Alzheimer's disease in the last 20 years.

To be fair, the company had plenty of help. Coauthors of their paper, "A mutation in APP protects against Alzheimer's disease and age-related cognitive decline" [1] include scientists and clinicians from Genentech in California; Landspítali University Hospital and the University of Iceland in Reykjavik, Iceland; The Institute for Human Genetics, Tubingen, Germany; Ulleva University Hospital and the University of Oslo, Norway; The Karolinska Institute and Hospital, Stockholm, Sweden; and The University of Helsinki, Finland. But the key to the study rests on two things: the use of genetically defined populations, and genome sequencing and analysis based on a specific hypothesis.

Before discussing what was found, it's useful to look at what's known about Alzheimer's disease and the problems associated with finding a treatment for it.

\*Correspondence: [petsko@brandeis.edu](mailto:petsko@brandeis.edu)  
Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA

Alzheimer's is the most prevalent form of dementia. About half of all people over the age of 85 develop it, but, although it is a disease of aging, it is not exclusively a disease of the aged: in the USA alone there are more than 5 million people suffering from it, and a considerable number are under 70 years of age. Auguste Deter, the first patient ever diagnosed by Alois Alzheimer (in 1901) as suffering from the disorder that now bears his name, was 51 when she was admitted with symptoms of dementia. By 2050, the number of Alzheimer's patients in the USA will exceed the populations of New York City, Los Angeles, Chicago and Houston combined. By the end of the century, it is estimated that there will be more than 300 million Alzheimer's sufferers worldwide and - since, on average, there are three care-givers for every patient - the disease will affect, directly or indirectly, a billion people. Alzheimer's disease is the only one of the ten leading causes of death today for which there is no treatment.

Among the known risk factors for Alzheimer's, the most dramatic is polymorphisms in the gene encoding apolipoprotein E, a protein that binds to a specific receptor on liver and peripheral cells. It is essential for the normal catabolism of triglyceride-rich lipoprotein constituents. Three alleles of this gene are associated with risk for dementia. The most common, *APOE3*, in homozygotes defines normal risk (that is, about 40% chance of developing the disease after age 85). The rarer *APOE2* and *APOE4* alleles are associated with reduced and elevated risk, respectively. E4/E3 heterozygotes have a significantly increased risk, while E4/E4 homozygotes have a greatly elevated risk: more than 80% of them may develop Alzheimer's by age 85. E2/E3 heterozygotes have lower incidence of Alzheimer's disease than normal, and the very rare E2/E2 homozygotes are largely free of the disease.

The histopathological hallmark of Alzheimer's is the presence in the brain of extracellular amyloid plaques and intraneuronal tangles. Both comprise misfolded proteins, with the plaques consisting largely of A $\beta$ , a peptide fragment of an integral membrane protein called APP, and the cytoplasmic tangles are made up chiefly of hyperphosphorylated forms of a microtubule-stabilizing protein called tau. Genetic and biochemical evidence suggests that tau is probably downstream of A $\beta$  in the etiology of the disease. The deCode work focused on APP and A $\beta$ , and so will we.

The amyloid hypothesis of Alzheimer's disease postulates that it is some species formed during the aggregation of the A $\beta$  peptide that initiates neuronal cell death. Therapeutic approaches in recent years have focused on getting rid of A $\beta$  (for example by plaque-clearing antibodies), or on preventing its formation in the first place.

A $\beta$  is produced by the (apparently) aberrant proteolysis of APP. Two cleavage events are required to make the A $\beta$  peptide. The first proteolytic enzyme is so-called  $\beta$ -secretase or BACE1, which cuts APP in a part of its sequence that lies outside the membrane. The second enzyme is a multiprotein complex called  $\gamma$ -secretase, which is a very unusual enzyme in that it cleaves APP *within* the membrane and is itself an integral membrane protein. Both BACE1 and  $\gamma$ -secretase have other functions in the cell, so it is not clear that inhibiting these proteases will produce a therapeutic benefit free from side-effects; nevertheless, inhibitors of both are currently in clinical trials for Alzheimer's disease.

All of which sounds sensible and logical except for the fact that there has been no genetic or other evidence to prove definitively that A $\beta$  production is a valid therapeutic target. The amyloid hypothesis, though resting on many pieces of evidence like this, is still a hypothesis. True, there are mutations in the genes encoding  $\gamma$ -secretase and APP that are associated with rare, familial forms of Alzheimer's, but animal models of Alzheimer's disease, created by overexpressing A $\beta$  in, for example, mouse brains, fail to mimic many of the details of the human disorder. No form of Alzheimer's is known that does not involve the tau protein as well as A $\beta$ , but there are other forms of dementia in which tau only is the causative agent, with no abnormal A $\beta$  production seen at all. It is not even clear that A $\beta$  itself is the initiating toxic agent: many people believe that one of the other fragments of APP produced during the aberrant proteolytic processing might actually be the pathological species. To make matters worse, even if A $\beta$  was the culprit, no one has any idea how much of it could be tolerated before the disease sets in or, to put it another way, how much the A $\beta$  level would need to be lowered, by antibody treatment or protease inhibition or some other mechanism, before a therapeutic benefit would be achieved. Recent clinical trials of several anti-A $\beta$  antibodies in patients with mild to moderate Alzheimer's disease have failed to meet their primary endpoints, though whether that is because the amyloid hypothesis is flawed or the treatment was administered much too late in the course of the disease remains unknown.

The new work from deCode provides what I believe are the first concrete answers to many of these questions. The data come first from studying a population of 1,736 Icelanders, and instead of doing the customary looking for families with a history of having Alzheimer's disease, the study's authors looked for families with a history of NOT having it. In particular, they looked for families with a low or zero incidence of the disease, even in members over 80, and hypothesized that such families might be protected by an allele in a known Alzheimer's-associated gene. For a variety of reasons it's harder to find

protective alleles than it is to find damaging ones, but if you can find both in the same gene they are among the most powerful evidence for a valid therapeutic target. And that is where deCode's genome mining operation struck gold, because the protective allele they found was in the *APP* gene, and its location was striking: a non-synonymous change to a threonine in the codon for alanine 673, a residue very close to the precise site where BACE1 cleaves APP to initiate the production of A $\beta$  (Ala673 is in fact the second residue in the sequence of the A $\beta$  peptide). Fascinatingly, this same residue has recently been found mutated to a valine in a single family with early-onset Alzheimer's disease, suggesting that the site is a critical one for the proteolytic processing of APP.

deCode and their collaborators then looked outside Iceland and found the same rare allele in some families elsewhere, chiefly in Scandinavia, and once again it was associated with very low risk of Alzheimer's disease. In fact it was found that the protective effects of A673T even extended to 25 homozygous carriers of the *APOE4* allele - something that would have been thought impossible by most Alzheimer's researchers.

That A673T is not common is unfortunate for most of us, to be sure, but it also means that the mutation could only be found by the kind of hypothesis-driven genomics that was undertaken. The expensive and, I would argue, disappointing genome-wide association studies so beloved by Big Science aficionados look for common variants that have small effects on risk, and so would never have found this one, just as they would never have identified the relatively uncommon *BRCA1* breast cancer mutations. Rare private variants are hard to identify, but when associated with disease they almost always have a relatively big effect, and therefore have a lot to teach us. Genome-wide association studies have largely taught us how to waste money.

Most of the people identified as carriers for A673T are heterozygotes, but a few elderly homozygotes were found, and none of them had any signs of dementia. Interestingly, the effect of the protective allele appears to include protection from age-associated mild cognitive impairment (you know, the more severe form of the "what was it I came down to the basement to find?" syndrome) as well as from dementia, and this effect is also stronger when using elderly controls than when using general population controls, presumably due to a greater frequency of the variant in the elderly. The study's authors estimate that the odds for carriers of the A673T allele of reaching age 85 are 1.47-fold that of non-carriers.

There are a number of ways the formation of an abnormal peptide could cause problems for a cell, many quite indirect - and there might, of course, be many other ways to trigger those same indirect processes. So the fact that A $\beta$  was associated with Alzheimer's disease didn't

really establish that it caused the disease. But if there are multiple pathways leading to Alzheimer's disease, it now seems, thanks to the deCode-led study, that they must go through A $\beta$ , because the effect of this protective allele is to reduce A $\beta$  formation, and people with both copies of the allele appear to be completely protected, or close to it. In human genetics, that is a good a smoking gun as you could ask for. I need no other: for me, this work proves that the production of the A $\beta$  peptide (note: not necessarily the formation of its amyloid plaques) is a necessary condition for the development of Alzheimer's disease, and therefore the processes that give rise to A $\beta$  are valid targets for treatment - or, more likely, prevention.

I say prevention because of two aspects of this terrific piece of genomics that have been almost completely ignored in the press reports and subsequent discussion. The first is that we now have an estimate of the amount that A $\beta$  needs to be lowered to have a therapeutic benefit. The authors of the study actually determined the production of A $\beta$  in 293T cells transfected with wild-type or mutant APP. The production of amyloidogenic peptides was 40% less by the A673T variant than by wild-type APP. For comparison, they also analyzed APP cleavage by the pathogenic A673V variant. In contrast to A673T, the A673V substitution resulted in markedly increased APP processing to A $\beta$ . (To confirm these *in vivo* observations, they also used an *in vitro* BACE1 cleavage assay to assess processing of a wild-type synthetic APP peptide substrate compared to a peptide bearing the A673T substitution. The A673T APP peptide was processed 50% less efficiently than the wild-type substrate, supporting the conclusion that it codes a sub-optimal BACE1 cleavage site.) Therefore, we now know that reducing the level of A $\beta$  by about half is sufficient to have a protective effect against dementia of the Alzheimer's type. Most of the clinical anti-A $\beta$  antibodies and secretase inhibitors produce a much greater reduction than that.

Unfortunately, we also know - and this is the other underreported aspect of the study - that this protective lowering works when the individuals have the allele for their entire lives. There is no evidence that it would work if they suddenly developed the mutation in, say their 60s, or even their 30s. As a result of this paper, I think we need to adjust our way of thinking about Alzheimer's disease. Therapeutic intervention after symptoms appear is likely to be years, probably decades, too late. A recent report in *The New England Journal of Medicine* sheds additional light on this point [2]. The authors used imaging and biomarker assessments in presymptomatic individuals with familial Alzheimer's disease mutations and asked how long before these individuals developed signs of dementia did they show abnormalities in APP proteolytic products. The study showed that the changes

in levels of A $\beta$  in the cerebrospinal fluid that accompany A $\beta$  deposition in the brain can be detected as long as 25 years before a particular person's symptoms begin. This implies that preventive treatment, at least in persons with mutations for autosomal dominant Alzheimer's disease, might need to begin 25 years, or even more, before conventional diagnosis of disease onset. In a perceptive commentary in the same issue, leading Alzheimer's researcher Dr Sam Gandy argues that we may need to think about treating some people for Alzheimer's disease throughout the course of their entire lives [3].

My assessment is the same as his. It may be that Alzheimer's disease needs to be thought of as similar to inborn errors of metabolism, which are treated by drugs or special diet almost from birth. Alternatively, we routinely vaccinate all children against measles; maybe we should also vaccinate them against A $\beta$ . Skeptics would argue that we do not know the potential deleterious effects of lifelong preventive therapy for Alzheimer's, and I guess that's true, but we do know that the A673T carriers live their entire lives with only about half the A $\beta$  of "normal" people - and not only do they seem fine, their life expectancy is actually longer.

Sam Gandy sums it up perfectly, "...reduction of the risk of late-life dementia requires a long-term and possibly lifelong effort" [3]. What is also clear - regardless

of whether, in light of the protective APP mutation, one considers the "amyloid hypothesis of Alzheimer's disease" as proven or not - is that any comprehensive strategy aimed at reduction of late-life dementia risk will almost certainly include monitoring and immunopharmacologic or neuropharmacologic control of A $\beta$  metabolism. deCode may never make enough of a profit to satisfy its investors, but, in terms of impactful science, they have just paid a huge return.

Published: 30 October 2012

#### References

1. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, Stefansson H, Sulem P, Gudbjartsson D, Maloney J, Hoyte K, Gustafson A, Liu Y, Lu Y, Bhangale T, Graham RR, Huttenlocher J, Bjornsdottir G, Andreassen OA, Jonsson EG, Palotie A, Behrens TW, Magnusson OT, Kong A, Thorsteinsdottir U, Watts RJ, Stefansson K: **A mutation in APP protects against Alzheimer's disease and age-related cognitive decline.** *Nature* 2012, **488**:96-99.
2. Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, Marcus DS, Cairns NJ, Xie X, Blazey TM, Holtzman DM, Santacruz A, Buckles V, Oliver A, Moulder K, Aisen PS, Ghetti B, Klunk WE, McDade E, Martins RN, Masters CL, Mayeux R, Ringman JM, Rossor MN, Schofield PR, Sperling RA, Salloway S, Morris JC; Dominantly Inherited Alzheimer Network: **Clinical and biomarker changes in dominantly inherited Alzheimer's disease.** *N Engl J Med* 2012, **367**:795-804.
3. Gandy S: **Lifelong management of amyloid-beta metabolism to prevent Alzheimer's disease.** *N Engl J Med* 2012, **367**:864-866.

doi:10.1186/gb-2012-13-10-176

**Cite this article as:** Petsko GA: Hypothesis-driven genomics pays off. *Genome Biology* 2012, **13**:176.