

Identifying Disease Genes – The First Step to Personalised Medicine

a report by

Jörg Hager

Chief Scientific Officer, IntegraGen SA

Introduction

Complex diseases such as obesity, Type 2 *diabetes mellitus* and many central nervous system (CNS) diseases affect millions of people worldwide. Yet, for many complex reasons, the therapeutic options available to tackle them successfully are limited. IntegraGen, a biotechnology company based in France, is applying its novel proprietary technology to map the genes associated with complex diseases and thereby provide new targets for drug development.

Limitations of Current Medicine

It is well recognised that, in many cases, only 30% to 40% of patients suffering with a disease respond positively to a particular drug, either because the drug given fails to target the right pathway for a specific patient or because the patient has a genetic mutation in one of the drug metabolising enzymes, reducing the effectiveness of the treatment. Most importantly, many diseases are diagnosed too late in life for drug treatment to be efficient because the disease has already progressed beyond the point of no return.

Early diagnosis, even before clinical onset, would be possible with greater knowledge of the genes causally implicated in each disease. Drugs could also be designed to target specific pathways implicated in the disease, resulting in higher positive response rates. Obviously, neither of these approaches is possible without knowledge of the genes and pathways involved and the appropriate molecular diagnostic tools.

Complex Genetics

A meaningful molecular diagnosis depends on the identification of the causal genes for the disease. Only this will provide sufficiently informative disease markers for a diagnostic test. However, for a complex disease like obesity, finding a single gene associated with the disease will be insufficient because it is likely that many disease-associated genes act synergistically to cause the condition. This has already been shown to be the case for some complex

diseases like Type 2 diabetes and schizophrenia. In schizophrenia, for example, one gene acting alone only increases the risk of disease by about 60%, whereas a combination of two genes increases the risk by up to 500%.

The most efficient way to identify genes that may act together in this way is to carry out genome-wide scans. Unfortunately, most technologies used today for genome-wide scanning are slow and very expensive, relying on time-consuming characterisation of many genetic markers (several hundreds in the case of microsatellite markers to several thousands in the case of single nucleotide polymorphisms (SNPs)). In addition, most technologies only allow the sequential genotyping of these markers and all rely on polymerase chain reaction (PCR) amplification, which is both error-prone and costly.

GenomeHIP®

IntegraGen's technology, called Genome Hybrid Identity Profiling (GenomeHIP®), offers a new approach. It identifies DNA regions that are linked to disease without cumbersome marker-by-marker genotyping at a high genome-wide resolution. Isolating disease-associated genes in this way gives researchers a fast, precise way of identifying disease markers and new potential drug targets for the treatment of complex diseases.

The approach taken by GenomeHIP® technology involves three critical steps:

- mixing genomic DNA from two related individuals with the same disease;
- the elimination of all DNA segments that are not identical between the individuals (as they share the disease, the genes for the disease will be within the identical segments of their genomes); and
- identification of the identical segments through hybridisation to a DNA microarray representing the human genome.

In this way, GenomeHIP® processes genomic DNA



from pairs of related individuals, such as sibling pairs. Comparison of many pairs allows researchers to identify a few target regions of the genome most likely to carry genes involved in the disease. This method generally requires much lower population sample sizes than the more traditional SNP or microsatellite studies, thanks to a higher degree of informativity, with an average of only 150 sibling pairs being needed for each study.

The Next Generation in Gene Discovery Technology

Developments in molecular biology techniques have already greatly increased our knowledge of diseases caused by single gene defects. The next challenge is to investigate the genetics of complex diseases such as obesity, *diabetes mellitus*, asthma and coronary artery disease. This research is made much more difficult by the fact that these diseases are likely to involve a combination of genetic factors, rather than just a single gene.

For genetic analysis of complex diseases to be practical, the methods involved must be fast, accurate, reliable and precise and have high statistical power. The size of the population required for familial studies is also a very important factor. To date, the two approaches of single nucleotide polymorphism analysis and microsatellite analysis have been used.

SNPs and microsatellites are types of genetic variation that occur at regular intervals and at high frequency throughout the human genome, so can be used as 'markers' within the genome. If a particular marker is found to be common amongst people with a particular disease, it will suggest that the gene involved is probably located near the marker.

SNPs are the most common type of DNA sequence variations found in humans. A SNP is where a single DNA base varies between different people and they occur once every 500 to 1,000 base pairs. To be classified as an SNP, a polymorphism must be present in at least one percent of the population. The even distribution of SNPs throughout the human genome, combined with their frequency and stability, make them valuable as genetic markers. Many genetic diseases are either caused by, or closely associated with, specific SNPs. SNP maps of the genome have already been used to identify genes involved in psoriasis, migraine, Alzheimer's disease and diabetes.

To date, microsatellite marker analysis has been the most successful tool for locating genes involved in single gene disorders, but they have also proved useful as markers for complex diseases. Microsatellite markers close to the disease gene correlate with the heredity of

the disease and, by analysis of these markers within families, researchers can predict how the disease will be inherited. Microsatellites do not occur as frequently as SNPs and because of this, the location of the relevant gene may not be identified as precisely as with other technologies. However, as microsatellite markers are more variable, they are usually much more informative than SNPs. Recently, researchers identified an area of the genome conferring susceptibility to both schizophrenia and bipolar disorder on chromosome 3, using 388 microsatellite markers within eight families. A similar technique has also been used to identify a putative asthma susceptibility gene.

Despite great improvements in SNP and microsatellite-based techniques, they still rely on the time-consuming and often sequential characterisation of many genetic markers. All methods involve PCR amplification of samples, which increases the costs significantly.

In comparison, GenomeHIP® is a one tube procedure that enables the quick comparison of the genomes of two related individuals who are afflicted with the same condition. Instead of characterising genetic markers, GenomeHIP® eliminates all the regions that differ between two entire genomes in a single experiment. The remaining identical genome segments will, by definition, include genes that are associated with the disease.

The major advantages of GenomeHIP® over marker-based approaches are:

- GenomeHIP® is based on the physical mapping of the genome;
- it identifies inherited sections of the genome without access to parental genomes;
- fewer affected individuals are needed;
- involves fewer false positives (more robust);
- higher resolution;
- increased statistical power; and
- there is no influence from sex-specific recombination differences.

Proven Effectiveness

IntegraGen has recently proven the power of this new technology when GenomeHIP® was used to map short regions of the genome likely to contain five genes associated with early onset obesity and six regions linked to autism. Both genome-wide studies were completed in just three months.

The average size of the identified regions was less than two megabases, considerably smaller than the typical size of the region identified by SNP or microsatellite analysis. The high resolution of GenomeHIP technology meant that only a couple of

genes had to be screened for each region identified, allowing the rapid discovery of three genes associated with obesity and, so far, one gene associated with autism. IntegraGen has a strong interest in these two diseases and plans to develop new diagnostic tests based on these results.

IntegraGen also successfully mapped a gene locus for mastocytosis in a collaborative project and, using just 23 sample pairs, the company identified an important region containing a strong candidate gene in less than two months. These examples show the power and superiority of IntegraGen's GenomeHIP® platform over traditional marker-based linkage mapping approaches.

The Future in Mapping the Genetic Causes of Disease

IntegraGen is focussing on four common complex diseases. In addition to the initial success seen in its investigation into the genetics of obesity and autism, the company also has projects running for Type 2 diabetes and bipolar disorder. All of these diseases have a high or increasing prevalence and, as such, are of great importance to society. IntegraGen has established a strong network of academic and clinical collaborations to support its investigations by providing access to valuable sample collections and disease-related expertise.

As has already been shown for obesity, the genes that IntegraGen discovers associated with these diseases will shed light on the metabolic pathways involved in the disorders and offer interesting targets for potential new drugs. IntegraGen also expect to be able to develop important diagnostic tests that will help in the diagnosis and appropriate treatment of the conditions.

IntegraGen's business strategy is to enter into partnerships with companies wishing to work with the company to develop the targets that IntegraGen identifies in these disease areas or work with them to identify new targets for other diseases, whilst retaining the rights to develop diagnostic tools. IntegraGen offers access to a unique, novel and rapid gene discovery technology, which will be valuable in the search of novel drug targets for complex diseases and the development of important diagnostic tools. ■

Contact Information

IntegraGen

4, rue Pierre Fontaine

91000 Evry

France

Tel: +33 (0) 1 60 91 09 00

Fax: +33 (0) 1 60 91 09 22

e-Mail: contact@integragen.com

<http://www.integragen.com>